

Sistema de Visão Stereo Ativo aplicado aos Robôs do ISePorto

Rui Miguel Santos Serra

1070308

— *orientador* —

José Miguel Soares de Almeida (Mestre)

ISEP/LSA

Tese submetida no âmbito do

Mestrado em Engenharia Eletrotécnica e de Computadores

Ramo de Sistemas Autónomos

NOVEMBRO DE 2012

MESTRADO EM ENGENHARIA ELETROTÉCNICA E DE COMPUTADORES

Sistema de Visão Stereo Ativo aplicado aos Robôs do ISePorto

por

Rui Miguel Santos Serra

LABORATÓRIO DE SISTEMAS AUTÓNOMOS

INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

Tese de mestrado orientada por:

José Miguel Soares de Almeida, *Mestre*

LABORATÓRIO DE SISTEMAS AUTÓNOMOS

INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

Porto, Novembro 2012

Resumo

A presente dissertação endereça o desenvolvimento de um sistema de visão *stereo* ativo para os robôs de futebol robótico da equipa ISePorto do ISEP, de modo a que estes tirem o máximo partido das câmaras rotativas neles existentes.

Este trabalho surge da necessidade de melhorar a capacidade de perceção do ambiente por parte dos robôs, principalmente da perceção da bola quando não está no plano do campo e dos robôs adversários. Esta necessidade surge devido ao aumento da dinâmica que se tem vindo a verificar ultimamente nas competições.

Para tal, foram estudados algumas trabalhos relacionados no que diz respeito a sistemas de visão *stereo* com *baselines* variáveis e eixos de rotação em ambas as câmaras, bem como fundamentos de visão *stereo*.

Foi proposta uma arquitetura para o sistema de visão ativo de modo a ser aplicado em qualquer robô da equipa MSL (*Middle Size League*).

Para tornar possível a implementação desta arquitetura foi desenvolvido um procedimento para a calibração e determinação em tempo real dos parâmetros extrínsecos do par *stereo* em função da posição angular dos eixos rotativos do robô. O sistema de visão foi também dotado de capacidade de sincronismo e foram implementadas funcionalidades ao nível de software que possibilitam a deteção de objetos na imagem, a correspondência de objetos presentes nas imagens de ambas as câmaras e consequentemente a determinação das posições tridimensionais desses objetos relativamente ao robô.

O sistema desenvolvido foi testado e validado em cenário MSL ao nível de perceção da bola, robôs adversários e linhas do campo.

Os resultados obtidos apresentam uma melhoria significativa, face à implementação atual dos robôs, na perceção tridimensional da bola quando não está no plano do campo, e dos robôs adversários.

Palavras-chave

Visão por Computador, Sistemas Autónomos, Robótica Móvel, Sistema de Visão *Stereo* Ativo.

Abstract

The current thesis addresses the development of an active stereo vision system for the Soccer robots of the IsePorto Team from ISEP, so they can get maximum advantage of their rotational cameras.

The need to improve the robot's surrounding environment perception and mostly the ball perception when it's not on the field plan and also, of the opponent players. This improvement on the robot perception becomes crucial with the latest increases of dynamic inside the MSL competition.

To do such task, it have been studied some active stereo vision techniques with variable baselines and rotation axes in both cameras.

It was proposed an active vision system architecture that can be applied to any MSL (Middle Size League) team robot.

To turn possible the architecture implementation, it were developed the necessary procedures to real time calibration and determination of the extrinsic parameters of the stereo couple according to angular position of the robot rotational axes. The vision system was also gifted of synchronism abilities and some software functions, which make possible, objects association within the images captured from both cameras and the consequent determination of the tridimensional positions of those objects relatively to the robot.

The system was tested and the results were validated, in a MSL scenario, regarding ball perception, opponent robots and field limit lines.

The obtained results show significant improvements relatively to current robot implementation. In the ball tridimensional perception when it's not on the field plan and the opponent robots.

Keywords

Computer vision, Autonomous Systems, Mobile Robotics, Active Stereo Vision System.

Agradecimentos

Agradeço ao Eng. José Miguel Almeida a sua orientação pela ajuda a progredir concetualmente na execução do trabalho desta dissertação. A todos os membros do LSA, em especial ao Eng. André Dias, pela disponibilidade, a amizade e contribuições na conclusão desta dissertação.

Aos meus pais, por sempre me terem ajudado e acarinhado no culminar de uma etapa fundamental da minha vida académica.

À Ana, pelo seu carinho, confiança e apoio constante que permitiram a concretização da minha dissertação.

Em memória da minha Avó...

A Todos Muito Obrigado,
Rui Miguel Santos Serra

Esta página foi intencionalmente deixada em branco.

“Se o produto escrito é honesto, não pode ser separado da pessoa que o escreveu.”

Tennessee Williams

Esta página foi intencionalmente deixada em branco.

Conteúdo

Resumo	v
Abstract	vii
1 Introdução	1
1.1 Âmbito da Dissertação	1
1.2 Enquadramento e Motivação	3
1.3 Objetivos	4
1.4 Estrutura da Dissertação	4
2 Estado da Arte	7
2.1 Sistemas Visão <i>Stereo</i>	7
2.1.1 Visão <i>Stereo</i> Ativa	8
3 Fundamentos de Visão	13
3.1 Introdução	13
3.2 Projeção Perspetiva	14
3.2.1 Parâmetros Intrínsecos	15
3.2.2 Parâmetros Extrínsecos	17
3.3 Visão <i>Stereo</i>	19
3.3.1 Geometria Epipolar	20
3.4 Calibração <i>stereo</i> através dos parâmetros extrínsecos	22
3.5 Retificação <i>Stereo</i>	22
3.6 Cálculo da triangulação <i>stereo</i>	24
4 Projeto	27
4.1 Formulação do Problema	27
4.2 Arquitetura do Sistema	30
4.3 Calibração do Sistema <i>Stereo</i> Ativo	31

5	Implementação	35
5.1	Implementação do sistema de aquisição	35
5.1.1	Câmaras CCD	35
5.1.2	Dispositivo de disparo das câmaras e leitura do ângulo de rotação	36
5.2	Implementação do Sistema de Visão <i>Stereo</i> Activo	37
5.2.1	Geometria epipolar	38
5.2.2	Parâmetros extrínsecos e calibração <i>stereo</i> ativa	38
5.2.3	Determinação de objetos nas imagens	39
5.2.4	Determinação do Ponto 3D no mundo	40
6	Resultados	41
6.1	Determinação da posição da bola	41
6.1.1	Posição da bola com a câmara <i>head e keeper</i>	42
6.1.2	Posição da bola com a câmara <i>head e kicker</i>	44
6.2	Determinação da distância aos robôs adversários	45
6.3	Determinação da distância as linhas de campo	46
6.4	Validação da visão <i>stereo</i> ativo	46
7	Conclusão e Trabalho Futuro	51
	Bibliografia	55

Lista de Figuras

1.1	Robô da equipa de futebol robótico ISePorto	3
2.1	Câmara <i>Kinect</i>	8
2.2	Visão <i>Stereo</i> Actica [1].	8
2.3	Sistema Experimental [1].	9
2.4	Sistema de visão estéreo com <i>baseline</i> variável [2].	9
2.5	Estimação da triangulação através do sistema de visão <i>stereo</i> [2].	10
2.6	<i>Baseline</i> paralela e convergente [3].	10
2.7	Separação da <i>baseline</i> [3].	11
3.1	Modelo de uma câmara <i>pinhole</i> [4].	14
3.2	Imagens de calibração dos parâmetros intrínsecos de uma câmara.	16
3.3	Representação dos três ângulos de rotação numa câmara.	17
3.4	Rotação em torno do eixo dos zz 's.	18
3.5	Rotação em torno do eixo dos xx 's.	18
3.6	Rotação em torno do eixo dos yy 's.	19
3.7	Exemplo de uma estrutura de Visão estéreo horizontal.	20
3.8	Representação da geometria epipolar [5].	20
3.9	Retificação de duas imagens <i>stereo</i> e representação das linhas epipolares horizontais colineares [6].	23
3.10	Geometria após retificação - planos da imagem paralelos e epipolos no infinito	24
3.11	Triangulação <i>stereo</i> [7]	25
4.1	<i>Middle Size League</i> - MSL.	28
4.2	Robô da equipa RFC Stuttgart.	29
4.3	Robô da equipa TechUnited.	29
4.4	Posição das câmaras no robô guarda-redes ISePorto.	29
4.5	Eixos de rotação no robô ISePorto [8].	30

4.6	Arquitetura sistema de visão e integração no robô IsePorto.	31
4.7	Vista do alvo pelas câmaras.	32
5.1	Câmara UeYE modelo UI-1225-LE [9].	36
5.2	Dispositivo de disparo das câmaras.	36
5.3	Encoder para determinação do ângulo da câmara superior.	37
5.4	Sistema Embebido Pandora Motion [8].	37
5.5	Deteção dos cantos encontrados no alvo.	37
5.6	Linhas Epipolares.	38
6.1	Determinação da posição da bola na situação 1.	42
6.2	Determinação da posição da bola na situação 2.	42
6.3	Determinação da posição 3D de uma bola aérea.	43
6.4	Determinação da posição da bola na situação 3.	44
6.5	Determinação da posição da bola na situação 4.	44
6.6	Determinação da posição do robô adversário na situação 1.	45
6.7	Determinação da posição do robô adversário na situação 2.	45
6.8	Determinação da distância do canto ao robô.	46
6.9	Posição da bola em X relativamente ao valor do eixo de rotação da câmara <i>head</i>	47
6.10	Posição da bola em Y relativamente ao valor do eixo de rotação da câmara <i>head</i>	47
6.11	Posição da bola em Z relativamente ao valor do eixo de rotação da câmara <i>head</i>	48
6.12	Erro da posição da bola em X relativamente ao valor da medida exata.	48
6.13	Erro da posição da bola em Z relativamente ao valor da medida exata.	49

Lista de Tabelas

6.1	Resultados obtidos da posição da bola entre a câmara <i>head</i> e <i>keeper</i>	42
6.2	Resultados obtidos da posição da bola aérea entre a câmara <i>head</i> e <i>keeper</i> . .	43
6.3	Resultados obtidos da posição da bola entre a câmara <i>head</i> e <i>kicker</i>	44
6.4	Comparação da distância ao robô adversário com duas <i>baselines</i> diferentes. .	45
6.5	Resultados obtidos da posição da bola entre a câmara <i>head</i> e <i>keeper</i>	46

Esta página foi intencionalmente deixada em branco.

Símbolos

X, Y, Z	-	Coordenadas de ponto no mundo
x, y	-	Coordenadas de ponto na imagem
x', y'	-	Coordenadas correspondente a x, y, z numa segunda imagem
R	-	Matriz de rotação
T	-	Vetor translação
m	-	Ponto na imagem
M	-	Ponto no mundo
m'	-	Ponto correspondente a m numa segunda imagem
f	-	Distância focal
P	-	Matriz projeção perspectiva
K	-	Matriz dos parâmetros intrínsecos
Π	-	Matriz dos parâmetros extrínsecos
I	-	Matriz identidade
e	-	Epipólo
u	-	Coordenada horizontal na imagem
l	-	Linha Epipolar
F	-	Matriz Fundamental
E	-	Matriz Essencial
ϕ	-	Ângulo de rotação <i>roll</i>
θ	-	Ângulo de rotação <i>pitch</i>
ψ	-	Ângulo de rotação <i>yaw</i>

Esta página foi intencionalmente deixada em branco.

Acrónimos

2D	-	Bidimensional
3D	-	Tridimensional
DSP	-	<i>Digital Signal Processor</i>
FPS	-	<i>Frames Per Second</i>
GPU	-	<i>Graphics Processing Unit</i>
ISEP	-	Instituto Superior de Engenharia do Porto
LSA	-	Laboratório de Sistemas Autónomos
MSL	-	<i>Middle Size League</i>
PPM	-	<i>Perspective Projection Matrix</i>
USB	-	<i>Universal Serial Bus</i>

Esta página foi intencionalmente deixada em branco.

Capítulo 1

Introdução

1.1	Âmbito da Dissertação	1
1.2	Enquadramento e Motivação	3
1.3	Objetivos	4
1.4	Estrutura da Dissertação	4

1.1 Âmbito da Dissertação

Esta dissertação aborda o desenvolvimento de um sistema de visão *stereo* ativo e a sua aplicação/validação em sistemas autónomos.

Nos dias de hoje os sistemas autónomos são utilizados por inúmeras áreas de intervenção humana desde aplicações industriais, em medicina, no ramo automóvel ou em monitorização ambiental.

Um sistema autónomo é um sistema onde não existe discordância entre a perceção sensorial adquirida pelo próprio e aquela que é necessária para a sua atuação, ou seja, um sistema autónomo necessita de informação adquirida pelo mesmo, por forma a possibilitar-lhe a tomada de decisões, estabelecer comportamentos e definir áreas de atuação [10].

Quando estamos perante um sistema autónomo, este é constituído por três componentes essenciais:

- Sensores
- Unidade processamento
- Atuadores

A presente dissertação incide na vertente sensorial, ou seja, através de um sistema de visão *stereo* o sistema autónomo pode determinar a posição 3D de um determinado objeto no mundo.

A Visão Computacional utiliza métodos de extração automática de informações úteis a partir de imagens, como por exemplo a realizada pelo ser humano no seu sistema visual. Estas imagens podem ser adquiridas por câmaras de vídeo ligadas a um computador [11].

Com a crescente evolução tecnológica, os dispositivos de aquisição de imagens, estão cada vez mais evoluídos a nível de resolução e taxas de aquisição de imagens, funcionalidades, aumento capacidades computacionais, sistemas embebidos com elevada capacidade computacional (GPU, DSP) e com preços cada vez mais acessíveis [12]. Com isto, a crescente aplicação de técnicas e métodos de Visão Computacional tem aumentado nas últimas décadas em diversas áreas, tais como:

- Medicina (análise da evolução de deformações, fabrico de próteses, simulações computacionais, ...)
- Indústria (controlo de qualidade, monitorização do funcionamento de máquinas, classificação de produtos pela sua dimensão, formato e cor, ...)
- Sistemas de navegação e de identificação (identificação de alvos, controlo de *robots*, ...)
- Sistemas de segurança (autenticação de pessoas por reconhecimento facial, monitorização de segurança através da deteção de movimentos, ...)

Com o decorrer da evolução, começaram a surgir novas técnicas que diretamente ou indiretamente, permitiram obter mais informações a partir de imagens, como o conhecimento das posições tridimensionais (3D) relativas a um determinado referencial ou orientações locais de superfícies adquiridas nas imagens.

Os sistemas de visão *stereo* nos dias de hoje são muito utilizados para reconstrução 3D de objetos, ou seja, estas soluções necessitam de obtenção de correspondências entre imagens de uma estrutura adquiridas segundo pontos de vista distintos. Após calibradas as câmaras e

respetivas correspondências é possível calcular por triangulação as coordenadas 3D do mundo associados aos pontos recolhidos por ambas as câmaras. Estas técnicas utilizadas necessitam de pelos menos duas imagens adquiridas de pontos de vista diferentes sendo a sua aquisição no mesmo instante, quando o ambiente a rececionar contem objetos em movimento.

1.2 Enquadramento e Motivação

Ao longo dos últimos anos, o Laboratório de Sistemas Autónomos (*LSA*) do Instituto Superior de Engenharia do Porto (*ISEP*), tem vindo a desenvolver diversos sistemas autónomos, a nível aéreo, aquático e terrestre para aplicações nos mais variados ambientes de operação. Surge assim, a necessidade de recorrer a sistemas de visão para os robôs possam analisar o meio que os rodeiam [13].

Os sistemas autónomos utilizados nesta dissertação foram os robôs da equipa de futebol robótico MSL IsePorto.



Figura 1.1: Robô da equipa de futebol robótico IsePorto

O futebol robótico é uma área de investigação que tem evoluído nos últimos anos nas áreas de investigação artificial, cooperação entre agentes e robótica móvel [14].

Esta competição consiste num jogo de futebol entre equipas de cinco robôs completamente autónomos, o que exige que os robôs sejam capazes de autonomamente reconhecer os seus jogadores, a bola, a baliza e outros elementos do campo. Existe uma série de aspetos que têm que ser previamente calibrados e parametrizados para que a performance seja a melhor possível a nível da visão, tração e comunicação.

Desde o início da equipa ISePorto, os robôs tem evoluído a um ritmo elevado, tanto a nível de software como a nível de hardware, tendo aumentando de performance nos sistemas de remate, visão, implementação de táticas, jogadas em equipa, entre outras...

Um dos problemas encontrados na performance da equipa ISePorto é a determinação da distância tridimensional dos objetos ao robô, isto é, devido ao grande número de eixos de rotação presentes em cada robô torna-se complicado em tempo real determinar com bastante precisão através de visão monocular a distância à bola, a linhas do campo e a objetos fora do plano do campo como por exemplo as bolas pelo ar. Surge assim, o objetivo sobre o qual incide a presente dissertação, o desenvolvimento de um sistema de visão *stereo* ativo utilizando as câmaras existentes em cada robô.

1.3 Objetivos

Nesta dissertação é abordado o problema de percepção, ou seja, a determinação da posição tridimensional relativa de objetos pelos robôs da equipa de futebol ISePorto, tais como bola, linhas de campo e robôs adversários.

Assim sendo os objetivos desta dissertação passam por:

- Implementar um módulo para determinação, dinâmica e em tempo real da calibração *stereo* em função das medidas de ângulos dos vários eixos.
- Desenvolver um sistema de visão *stereo* ativo, isto é, que suporte pares de câmaras com movimento relativo entre si, que permita determinar a distância tridimensional aos objetos.
- Implementar e validar o sistema de visão *stereo* ativo na equipa ISePorto aplicando um mecanismo de disparo para a aquisição sincronizada das imagens com a leitura do ângulo do eixo de rotação que possibilitem a deteção de objetos na imagem, a correspondência de objetos presentes nas imagens de ambas as câmaras e consequentemente a determinação das posições tridimensionais desses objetos relativamente ao robô.

1.4 Estrutura da Dissertação

Esta dissertação encontra-se organizada da seguinte forma:

O primeiro capítulo apresenta a motivação que levou à elaboração desta dissertação, assim como a abordagem utilizada.

No segundo capítulo apresenta uma pesquisa sobre as abordagens utilizadas até ao momento sobre sistemas de visão *stereo* ativa, desde logo com o intuito de perceber que técnicas são utilizadas para calibração do sistema de visão.

De seguida, no terceiro capítulo, são expostos os conceitos e fundamentos de visão computacional necessários sobre projeção perspectiva e rotações 3D aplicada a visão para cálculo do sistema *stereo*.

No capítulo quatro é apresentado e analisado o problema e projetado sistema de visão *stereo* ativo.

Seguidamente, no quinto capítulo é apresentada a implementação do sistema de visão *stereo* ativo, a calibração do sistema e integração do sistema com ângulo de rotação.

No sexto capítulo, estão descritos os resultados obtidos e validação do sistema de visão *stereo* ativo.

Por fim, no sétimo capítulo são apresentadas algumas conclusões finais, bem como perspectiva de trabalho futuro.

Esta página foi intencionalmente deixada em branco.

Capítulo 2

Estado da Arte

2.1	Sistemas Visão <i>Stereo</i>	7
2.1.1	Visão <i>Stereo</i> Ativa	8

A pesquisa e conhecimento prévio, de um determinado tema, é essencial para a sua melhor compreensão e para a escolha da abordagem mais correta. Deste modo, é essencial o conhecimento prévio dos sistemas de visão *stereo* existentes.

2.1 Sistemas Visão *Stereo*

A determinação da posição tridimensional relativa aos objetos do meio envolvente em tempo real é um dos problemas abordados a nível de percepção em robótica.

Existem inúmeras aplicações que recorrem a sistemas de visão *stereo* para esse fim, utilizando soluções quer com *baselines* fixas, quer com *baselines* variáveis, isto é, a posição relativa das câmaras é variável, e normalmente controlada. Estes últimos sistemas são denominados de sistemas de visão *stereo* ativo.

Uma das câmaras utilizadas para determinação tridimensional é a *Kinect* (figura 2.1). Esta câmara utiliza um emissor de infravermelhos que projeta uma rede infravermelha em toda a superfície à sua volta e que se deforma formando uma rede com a forma da superfície em redor. No seu interior existe uma câmara de profundidade que analisa a malha de infravermelhos para construir um mapa em 3D do local e das pessoas que ali se encontram. É constituída por um motor TILT (movimento vertical) que ajusta a câmara para deteção de pontos de interesse.

A *Kinect* possui uma câmara normal que captura a informação em formato de vídeo para complementar a leitura dos objetos existentes. Toda esta informação é transmitida à unidade de processamento via USB. Contudo, existem vantagens e desvantagens nesta câmara. As desvantagens prendem-se pelo ângulo de visão da câmara ser muito reduzido e pouca resolução, contudo o custo torna uma grande vantagem na utilização desta câmara.



Figura 2.1: Câmara *Kinect*

2.1.1 Visão *Stereo* Ativa

A visão *stereo* analisa a informação tridimensional de objetos a partir de duas imagens capturadas no mesmo instante por duas câmaras colocadas no mesmo robô.

As técnicas de controlo de posição e orientação das câmaras num sistema de visão *stereo* ativo podem ser usadas para a correspondência de diferentes pontos vistos pelas câmaras para recuperar a profundidade e forma da superfície [1][15].



Figura 2.2: Visão *Stereo* Activa [1].

Na figura 2.2 podemos observar o robô de visão *stereo* ativo que consiste num sistema

binocular com movimento PAN (de rotação no plano) [15], sendo o controlo dos motores efetuado por sistema de controlo que analisa a posição dos três eixos de rotação através de *encoders*. A arquitetura do sistema em questão pode ser analisado na figura 2.3.

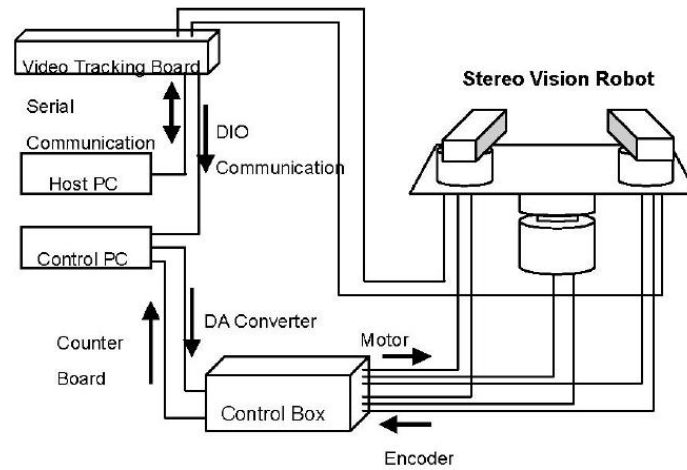


Figura 2.3: Sistema Experimental [1].

Na figura 2.4, foi desenvolvido uma *baseline stereo* variável utilizando um sistema com guias lineares com um deslocamento de alta velocidade (1ms) [2]. Para detetar um objeto em movimento, o sistema de visão captura uma imagem de cada câmara e estima a distância em apenas 1ms. Por sua vez, o cálculo para estimar a forma 3D do objeto demora cerca de 2ms.

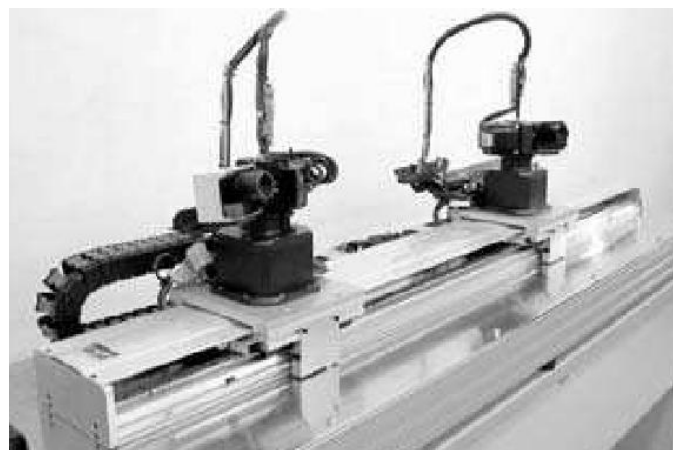


Figura 2.4: Sistema de visão estéreo com *baseline* variável [2].

Com a variação do comprimento da *baseline* como é possível observar na Figura 2.5, é possível melhorar a qualidade dos resultados obtidos da estimação 3D do objeto, eliminando assim erros de profundidade podendo rastrear um objeto mais rapidamente e com melhor precisão.

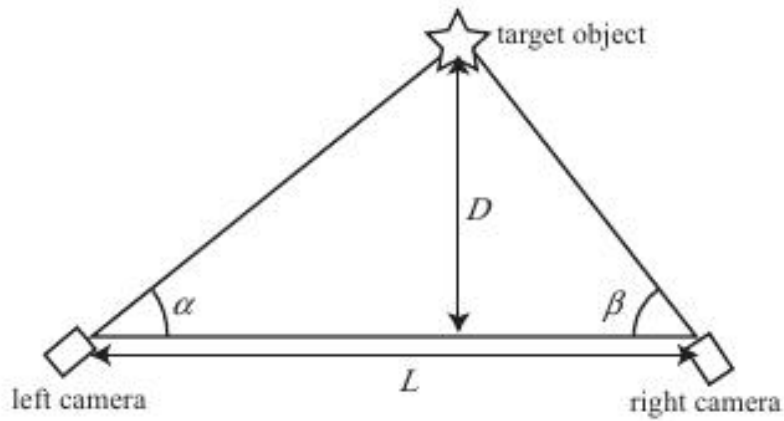


Figura 2.5: Estimação da triangulação através do sistema de visão *stereo* [2].

Em [3], foram utilizadas técnicas de visão *stereo* activas para controlar a posição e orientação das câmaras, usando a correspondência entre pontos de visões de câmaras diferentes para recuperar a profundidade da superfície e da sua forma, como se pode observar na figura 2.6.

Foram utilizados dois tipos de configuração de visão *stereo* como podemos observar na figura 2.6. A *baseline* paralela, oferece uma simples correspondência entre dois pontos vistos por cada câmara, contudo, a resolução da profundidade é baixa comparativamente com a geometria de imagens convergentes.

Para o bom funcionamento da visão *stereo* com câmaras convergentes, estas devem estar totalmente alinhadas relativamente a um ponto da imagem, ou seja, a distância focal de cada câmara a esse ponto deve ser igual [3].

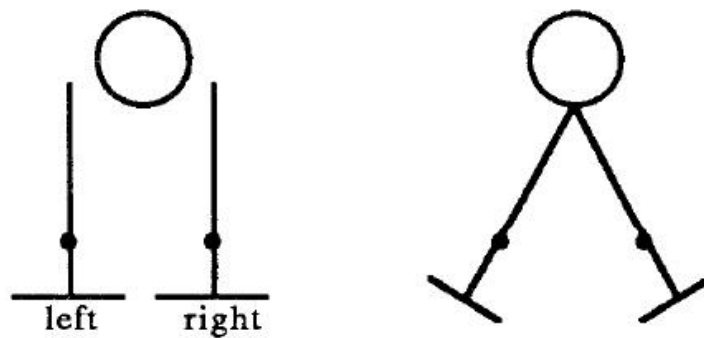


Figura 2.6: *Baseline* paralela e convergente [3].

Na figura 2.7 podemos visualizar o ponto do mundo $P(x, y, z)$ projetado segundo duas *baselines* diferentes, isto é, para uma *baseline* ($B1$) a diferença de projecção entre x' e x'' é

pequena, tornando o intervalo de *pixels* que devem ser pesquisados para estabelecer a correspondência pequenos.

Relativamente a uma *baseline* ($B2$), requer uma grande variedade de *pixels* a serem pesquisados, a fim de estabelecer a correspondência entre x' e x''' , fornecendo uma resolução de profundidade significativamente maior.

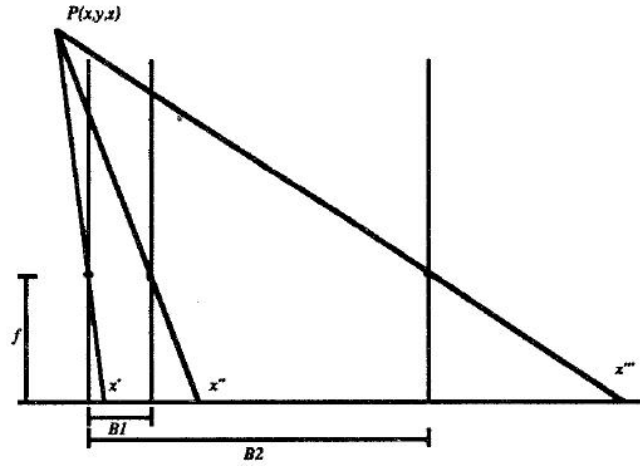


Figura 2.7: Separação da *baseline* [3].

Esta página foi intencionalmente deixada em branco.

Capítulo 3

Fundamentos de Visão

3.1	Introdução	13
3.2	Projeção Perspetiva	14
3.2.1	Parâmetros Intrínsecos	15
3.2.2	Parâmetros Extrínsecos	17
3.3	Visão <i>Stereo</i>	19
3.3.1	Geometria Epipolar	20
3.4	Calibração <i>stereo</i> através dos parâmetros extrínsecos	22
3.5	Retificação <i>Stereo</i>	22
3.6	Cálculo da triangulação <i>stereo</i>	24

3.1 Introdução

Visão *stereo*, ou *stereopsis*, designa *opsis* visão. O termo *stereo* é utilizado quando nos referimos à existência de mais do que uma imagem de uma determinada cena.

O ser humano tem a percepção tridimensional do que rodeia à sua volta, para isso utiliza duas imagens de um mesmo objeto capturada no mesmo instante pelos seus olhos, combinando-as de forma a obter uma percepção do que o rodeia de forma tridimensional.

Tal como o ser humano consegue estimar distâncias aos objetos utilizando os dois olhos, através de duas câmaras podemos determinar a distância a determinados objetos através de um sistema de visão *stereo*.

3.2 Projeção Perspetiva

Em visão por computador, o modelo mais comum de uma câmara é o *pinhole*. O modelo de câmara *pinhole* é um modelo matemático de uma câmara em que a lente é aproximada por um ponto no espaço, a uma distância do plano de imagem, (distância focal) por onde passa a luz. A projeção de pontos 3D em pontos do plano da imagem é uma projeção perspetiva, que pode ser descrita de forma elegante recorrendo à Geometria Perspetiva envolvendo coordenadas homogêneas [16] [17].

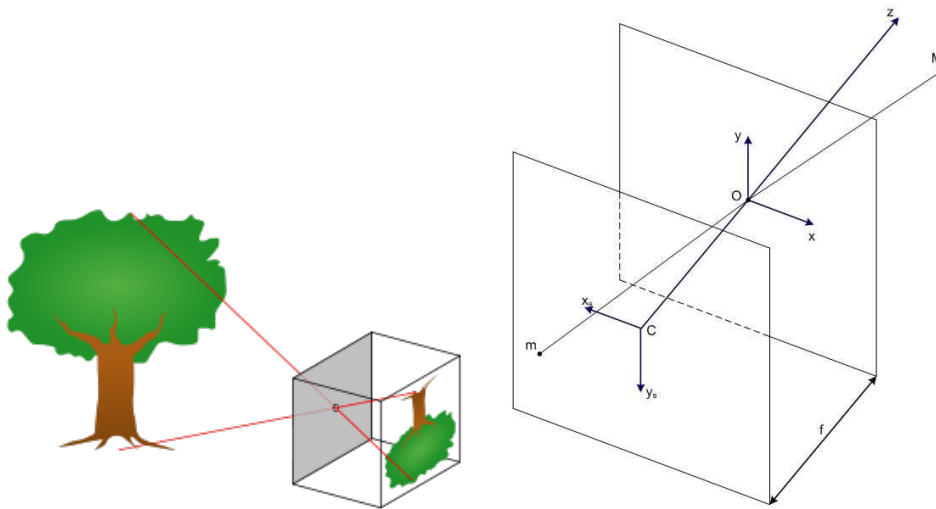


Figura 3.1: Modelo de uma câmara *pinhole* [4].

A imagem tridimensional, M , sofre uma projeção perspetiva, através da passagem pelo centro óptico, O , e consequente projeção no plano da imagem. O eixo óptico, é a linha que liga o centro óptico e o plano da imagem. A distância entre o centro óptico e o plano da imagem é a distância focal, f .

A projeção perspetiva pode ser obtida através de uma equação linear (3.1), o que torna o modelo de fácil utilização.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (3.1)$$

Em coordenadas homogêneas esta relação pode ser obtida da seguinte forma:

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = PPM \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.2)$$

onde PPM é a matriz de projeção perspectiva e que pode ser representa por:

$$PPM = A \begin{bmatrix} R|T \end{bmatrix} \quad (3.3)$$

Através da matriz PPM podemos projetar um ponto 3D do mundo num ponto de duas dimensões da imagem, onde a matriz A juntamente com os parâmetros de distorção, constituem os parâmetros intrínsecos da câmara, R e t são as componentes dos parâmetros extrínsecos e M um ponto do mundo.

$$m \approx (PPM)M \quad (3.4)$$

3.2.1 Parâmetros Intrínsecos

Os parâmetros intrínsecos servem para se poder relacionar as coordenadas *pixel* relativas às imagens com as coordenadas de pontos do espaço medidos no sistema referencial com origem no centro da câmara, ou seja, transforma as coordenadas 3D no sistema câmara em coordenada 2D no sistema da memória *frame*.

Estes parâmetros dependem exclusivamente das características físicas da câmara (da sua geometria interna e do tipo de lente) [5], isto é, inclui a informação associada à projeção perspectiva e resolução, alinhamento do sensor da imagem e num outro conjunto de coeficientes associados à distorção na imagem, provocada pela facto da lente não ser perfeita. Estes parâmetros estão presentes na matriz A da equação (3.5).

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.5)$$

Para obter os parâmetros intrínsecos e extrínsecos das câmaras utilizadas no sistema de visão, pode-se recorrer à utilização da ferramenta *Camera Calibration Toolbox* para Matlab [18].

O processo para obter os parâmetros intrínsecos, consiste em adquirir imagens de um alvo com um padrão de dimensões conhecidas, em diferentes configurações no espaço. Depois recorrendo a processos de otimização, é possível determinar o conjunto de parâmetros intrínsecos que otimizem o erro entre a projeção dos pontos 3D do alvo nas imagens e os correspondentes pontos detetados nas imagens. Posteriormente estes parâmetros são utilizados para o cálculo

dos parâmetros extrínsecos [19].

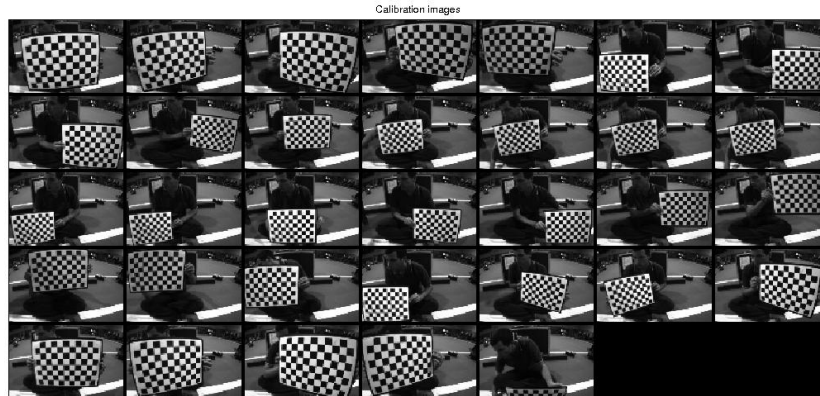


Figura 3.2: Imagens de calibração dos parâmetros intrínsecos de uma câmara.

Existem cinco tipos de parâmetros intrínsecos obtidos após a calibração:

- **Distância Focal:** representa a distância em *pixels*, entre o centro de projeção e o plano de imagem (segundo os seus dois eixos, x e y) e é guardada no vetor **fc** (2x1) (**fc(1)**= fc_x e **fc(2)**= fc_y). A distância focal efetiva (f) pode ser obtida a partir de: $fc_x \cong f \cdot s_x/d_x$ ou $fc_y \cong f/d_y$, onde d_x e d_y são as distâncias entre centros dos elementos sensores vizinhos segundo as direções X e Y , e s_x é um fator de incerteza horizontal devido a erros de sincronização;
- **Centro Ótico da Imagem:** representa as coordenadas, em *pixels*, do centro ótico da imagem na memória *frame* e são guardadas no vetor **cc** (2x1) (**cc(1)**= cc_x e **cc(2)**= cc_y);
- **Coefficiente de Desfasamento:** representa o ângulo entre os eixos x e y da imagem na memória *frame* e é guardado na variável **alpha_c**;
- **Distorções radial e tangencial:** representa os coeficientes de distorção da imagem (radial e tangencial) e são guardadas no vetor (5x1)**kc** (**kc(1)**, **kc(3)** e **kc(5)** são os coeficientes da distorção radial de 2^a, 4^a e 6^a ordem, respetivamente, e **kc(2)** e **kc(4)** são os coeficientes da distorção tangencial de 1^a e 2^a ordem, respetivamente).

3.2.2 Parâmetros Extrínsecos

Os parâmetros extrínsecos são utilizados para transformar as coordenadas 3D no sistema de coordenadas do mundo em coordenadas 3D no sistema da câmara, para cada posição e/ou orientação da câmara/plano de calibração [20]. Assim, existem seis parâmetros: os ângulos de rotação ϕ , θ e ψ , e as três componentes do vetor de translação T . A matriz de rotação R pode ser representada em função de ϕ , θ e ψ , como:

$$R = R_{(x,\phi)} R_{(y,\theta)} R_{(z,\psi)} \quad (3.6)$$

Uma rotação pura significa que dois dos eixos são fixos e o movimento rotacional é feito em volta do restante eixo [21]. O processo de rotação pura permite obter os três ângulos de rotação: *roll*, *pitch* e *yaw*, também conhecido por ângulos de *Euler*.

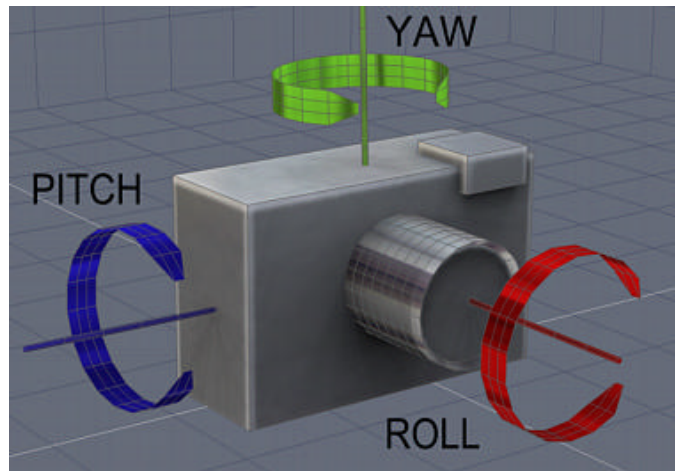


Figura 3.3: Representação dos três ângulos de rotação numa câmara.

Com isto, os três eixos base podem ser definidos da seguinte forma: rotação em *roll* (eixo dos zz 's), rotação em *pitch* (eixo dos xx 's) e rotação em *yaw* (eixo dos yy 's).

A rotação em *roll* é feita através da rotação dos eixos xx 's e yy 's em torno do eixo dos zz 's. Esta é descrita algebricamente pela seguinte relação:

$$R_{(z,\psi)} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

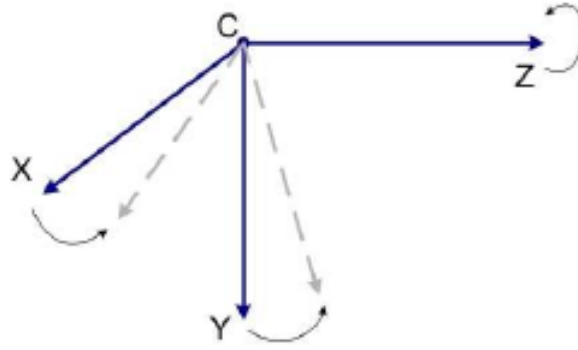


Figura 3.4: Rotação em torno do eixo dos zz 's.

A rotação em torno do eixo dos xx 's designada por *pitch*, é descrita algebricamente por:

$$R_{(x,\phi)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \quad (3.8)$$

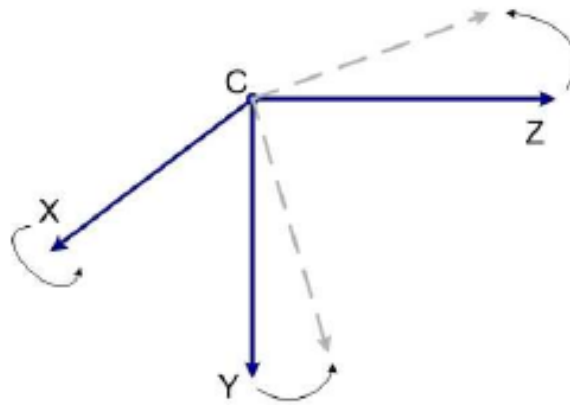
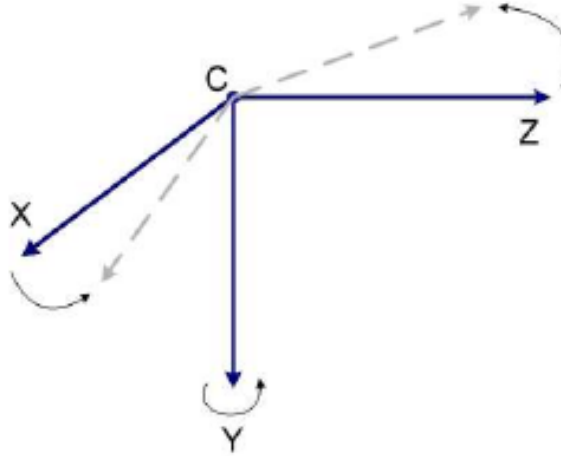


Figura 3.5: Rotação em torno do eixo dos xx 's.

Por fim, a rotação no eixo dos yy 's designada por *yaw*, matricialmente representa-se por:

$$R_{(y,\theta)} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \quad (3.9)$$

Figura 3.6: Rotação em torno do eixo dos yy 's.

A matriz de rotação R em função de ϕ , θ e ψ é representada por:

$$R = R_{(x,\phi)} R_{(y,\theta)} R_{(z,\psi)} = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ \cos\psi\sin\phi\sin\theta - \cos\theta\sin\psi & \sin\psi\sin\phi\sin\theta + \cos\phi\cos\psi & \sin\phi\cos\theta \\ \cos\psi\sin\theta\cos\phi + \sin\phi\sin\psi & -\sin\phi\cos\psi + \sin\psi\sin\theta\cos\theta & \cos\theta\cos\phi \end{bmatrix} \quad (3.10)$$

3.3 Visão Stereo

No sistema de visual humano, a percepção da profundidade é conseguida pela diferença entre a projeção dos pontos de uma estrutura 3D no olho esquerdo e direito. É a partir desta *visão estéreo* que o ser humano avalia profundidades ou julga distâncias.

Contudo, a perda de informação que ocorre ao projetar-se uma estrutura do mundo real 3D no plano bidimensional de uma imagem, faz com que geralmente haja ambiguidade na solução do problema da sua reconstrução ou que o referido problema seja mesmo impossível de resolver [22]. Como consequência, as técnicas de reconstrução 3D necessitam de impor certas restrições sobre o mundo real de forma a eliminar algumas ambiguidades na sua reconstrução. Estas suposições, de acordo com o número de vistas usadas na análise da estrutura a reconstruir, podem-se basear em técnicas monoculares, binoculares ou de múltiplas vistas [23].

As técnicas de visão *stereo* são a forma mais usual para se determinar as coordenadas dos pontos 3D de uma estrutura. A determinação de informação tridimensional 3D de pontos, usando apenas câmaras de imagem, envolve a utilização de pelo menos duas imagens, adquiridas sob pontos de vista distintos. Uma das formas para a aquisição das imagens da estrutura

em análise, consiste em afastar as duas câmaras de uma certa distância, à qual se chama linha de base (*baseline*), medida entre os seus centros óticos (Figura 3.7).

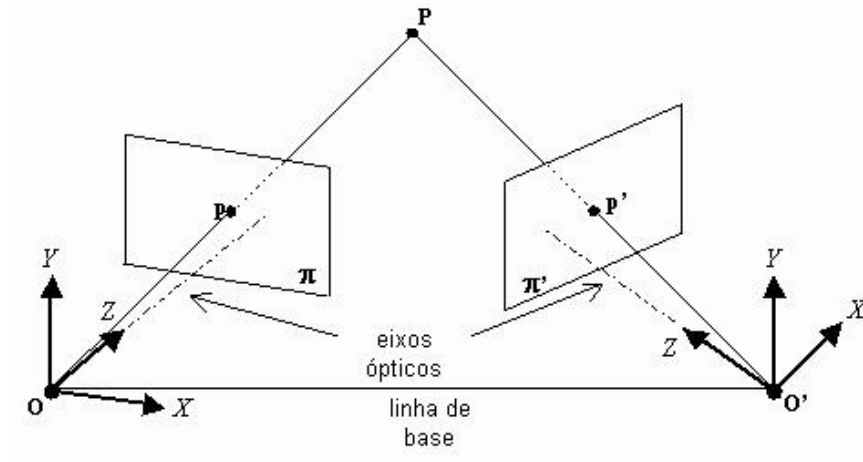


Figura 3.7: Exemplo de uma estrutura de Visão estereó horizontal.

Em sistemas *stereo* a correspondência entre imagens é normalmente sustentada nas restrições impostas pela geometria epipolar.

3.3.1 Geometria Epipolar

A geometria epipolar existente entre duas imagens é essencialmente a geometria da interseção dos planos de imagem com a *baseline* [5]. Na figura 3.7, está ilustrada o princípio da geometria epipolar *stereo*, onde o ponto tridimensional P é projetado no plano da imagem de duas câmaras, entre os quais se conhece a posição relativa (t) e a orientação relativa (R).

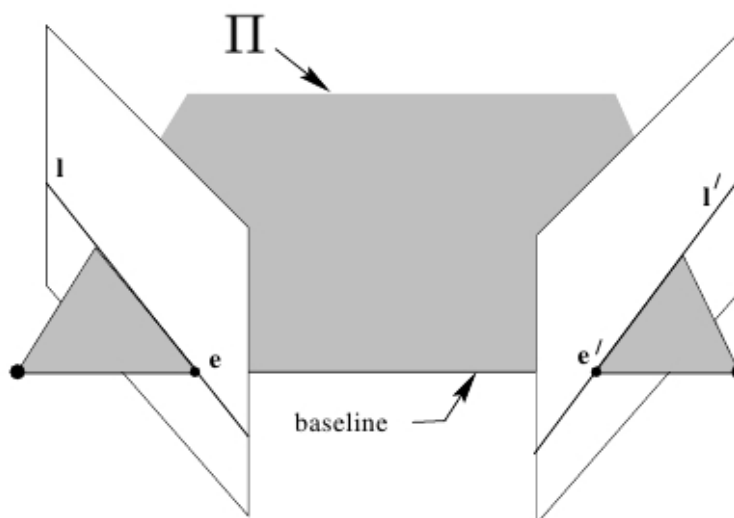


Figura 3.8: Representação da geometria epipolar [5].

Na figura 3.8 estão representadas as componentes da geometria epipolar. A *baseline* é

o eixo que une o centro das duas câmaras. Os epipólos \mathbf{e} e \mathbf{e}' são os pontos de interseção entre a *baseline* e os planos de imagem. O plano epipolar (plano Π) é o plano que contém a *baseline* e um ponto no mundo (X). As linhas epipolares (l e l') são a interseção dos planos epipolares com os planos da imagem. Há duas formas de representar a geometria epipolar algebricamente, ou seja, através da matriz fundamental ou através da matriz essencial que representa um caso particular da matriz fundamental.

A geometria epipolar *stereo* é definida para um par de câmaras com disposição relativa constante e é independente da calibração de ambas. Pelo facto de não incorporar os parâmetros intrínsecos, a geometria epipolar, expressa pela fórmula (3.11), estabelece relações entre os sistemas referenciais métricos das imagens.

$$p_r^T E p_l = 0 \quad (3.11)$$

Onde, p_r e p_l são os pontos da câmara direita e esquerda em coordenadas homogêneas, e a matriz E , denominada de matriz Essencial, pode ser obtida em função da rotação e translação entre os referenciais das duas câmaras, R e T , da seguinte forma:

$$E = \hat{T}R \quad (3.12)$$

Utilizando a matriz essencial é possível calcular a posição dos epipolos (3.13), bem como da linha epipolar numa imagem associada a um ponto da outra imagem (3.14).

$$e'^T E = 0 \quad (3.13)$$

$$E e = 0$$

Sendo l' e l as linhas epipolares, geradas pelos pontos p_l e p_r das outras câmaras, podem ser definidas, por:

$$l' \sim E p_l \quad (3.14)$$

$$l \sim E' p_r$$

Sendo l e l' no fundo vetores perpendiculares ao plano epipolar Pi a determinação da matriz fundamental é obtida de forma idêntica, mas como para este caso os pontos têm de se encontrar em coordenadas da imagem é necessário aplicar o inverso da matriz dos parâmetros intrínsecos aos pontos de forma a que estes fiquem em coordenadas normalizadas ($A l^{-1} m$ e

$Ar^{-1}m')$.

$$(Ar^{-1}m')^T E(Al^{-1}m) = 0 \quad (3.15)$$

Assim a matriz fundamental é determinada através da seguinte relação:

$$F = (Ar^{-1})^T E(Al^{-1}) \quad (3.16)$$

$$m'^T F m = 0$$

O baixo número de graus de liberdade da matriz essencial provoca uma ambiguidade de escala, contudo, dispõe de cinco graus de liberdade, para a matriz de rotação e translação [24]. Esta característica permite que a matriz essencial seja decomposta em rotação e translação a menos de um fator de escala.

3.4 Calibração *stereo* através dos parâmetros extrínsecos

A calibração *stereo* pode ser determinada através da calibração dos parâmetros extrínsecos de cada câmara, ou seja, se a matriz de rotação e translação da câmara esquerda (R_l e T_l) e a matriz de rotação e translação da câmara direita (R_r e T_r) estão no mesmo referencial do mundo, a matriz de rotação e translação (R e T) [7] do sistema *stereo* pode ser determinada pela seguinte fórmula (3.17).

$$R = R_r R_l^T \quad (3.17)$$

$$T = T_l - R^T T_r$$

3.5 Retificação *Stereo*

Uma das maneiras para podermos calcular a distância a um objeto, é através da retificação das imagens [25] [26]. O processo de retificação faz uma re-disposição dos pixels em cada imagem de forma a que o alinhamento dos centros focais e planos da imagem das duas câmaras sejam os mesmos. Com isto, os epipolos são projetados para o infinito, o que torna as linhas epipolares horizontais e colineares, facilitando a correspondência entre imagens. O processo utilizado para a deformação das imagens é baseado na aplicação da matriz homografia a cada imagem [27]. Contudo, os epipolos são colocados no infinito e as linhas epipolares tornam-se paralelas relativamente ao eixo de coordenadas u da imagem. Os métodos par cálculo das

matrizes de homografia partilham a principal preocupação de minimização da distorção a aplicar às imagens.

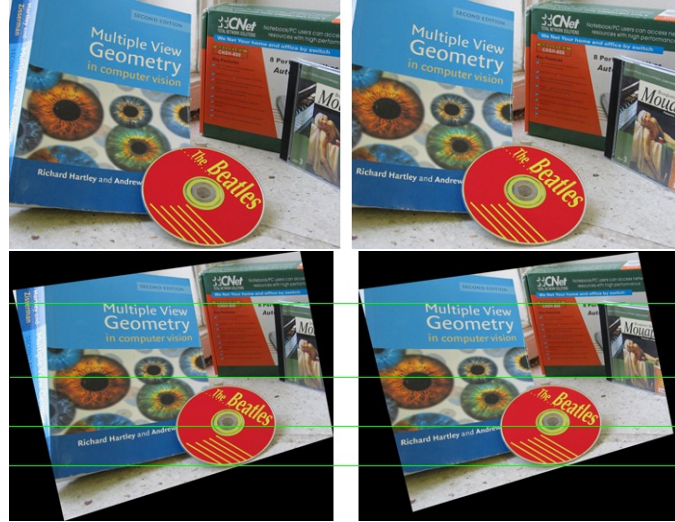


Figura 3.9: Retificação de duas imagens *stereo* e representação das linhas epipolares horizontais colineares [6].

Terminado o processo de retificação, a geometria do sistema *stereo* torna-se idêntica à apresentada na figura (3.10), com isto, o cálculo das coordenadas tridimensionais de pontos do mundo pode ser obtido pelas seguintes fórmulas (3.18, 3.19, 3.20).

$$x = \frac{x_l z}{f} \quad (3.18)$$

$$y = \frac{y_l z}{f} \quad (3.19)$$

$$z = \frac{fb}{x_l - x_r} \quad (3.20)$$

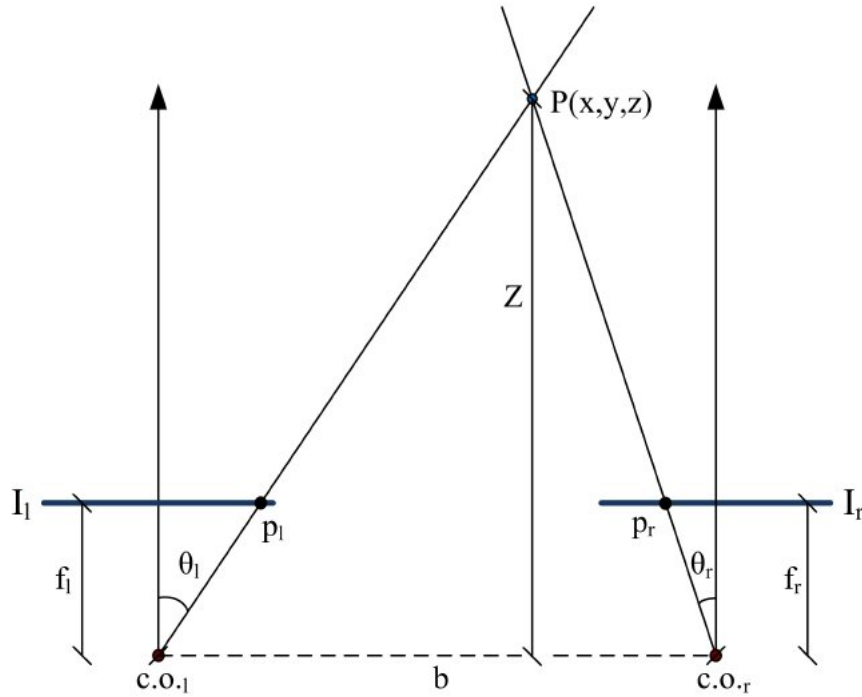
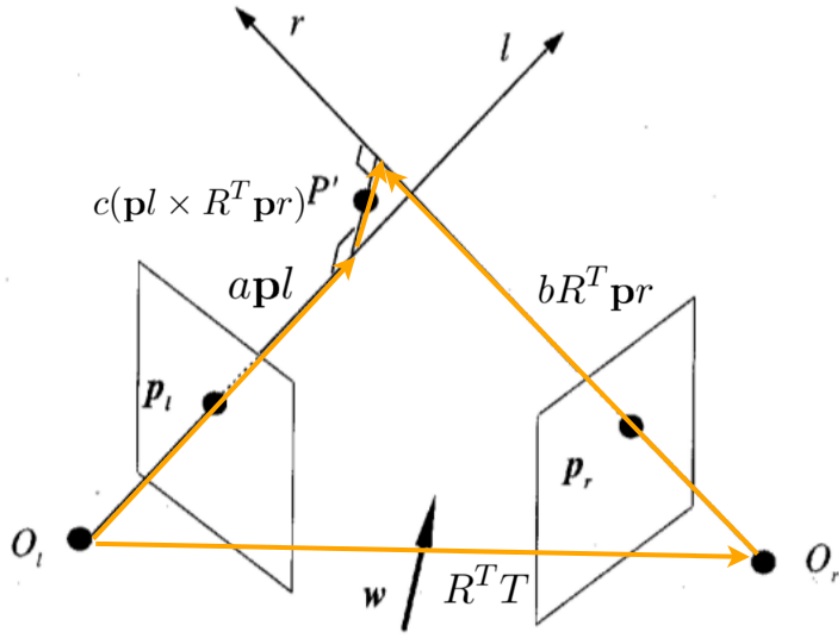


Figura 3.10: Geometria após retificação - planos da imagem paralelos e epipolos no infinito

Contudo como já mencionado anteriormente, o objetivo principal da retificação *stereo* é alinhar as duas imagens, uma de cada câmara, no mesmo plano da imagem, de maneira a que cada linha de pixels de uma imagem esteja alinhada com a linha de pixels da segunda imagem.

3.6 Cálculo da triangulação *stereo*

O cálculo da triangulação *stereo* permite determinar as coordenadas 3D sem recorrer à retificação *stereo* de um Ponto no mundo, dadas as suas projeções nas imagens adquiridas pelas duas câmaras. A fim de resolver este problema, é necessário conhecer os parâmetros do sistema de visão *stereo*, isto é, a matriz de Rotação (R) e a matriz de translação (T).

Figura 3.11: Triangulação *stereo* [7]

A translação(T), é obtida através da fórmula (3.21), isto é, T é o valor das coordenadas do centro ótico direito (O_r) menos as coordenadas do centro ótico esquerdo (O_l), no referencial da câmara direita.

$$T = O_r - O_l \quad (3.21)$$

Logo, o vetor T , no referencial esquerdo é definido por:

$$T = R^T T \quad (3.22)$$

Dado que se definiu que:

$$\mathbf{p}r = R\mathbf{p}l + T \quad (3.23)$$

Na equação (3.24), após conhecidas as matrizes de rotação(R) e translação(T) bem como os pontos da projeção da câmara esquerda ($\mathbf{p}l$) e direita ($\mathbf{p}r$) [7], podemos determinar o valor escalar (a , b e c), sendo a a distância segundo $\mathbf{p}l$, b a distância segundo $\mathbf{p}r$ (rodado para o referencial esquerdo), ou seja, $R^T \mathbf{p}r$ e c a distância segundo a direção perpendicular às retas $\mathbf{p}l$ e $\mathbf{p}r$.

$$a\mathbf{p}l - bR^T \mathbf{p}r + c(\mathbf{p}l \times R^T \mathbf{p}r) = R^T T \quad (3.24)$$

Após determinado o valor de a , b , e c , que verifica a equação (3.24), o ponto da interseção das duas retas (P'), pode ser definido pela seguinte equação (3.25).

$$P' = a\mathbf{pl} + \frac{c}{2}(\mathbf{pl} \times R^T \mathbf{pr}) \quad (3.25)$$

O valor escalar c fornece a medida da distância entre as duas linhas de visão, obtendo assim, o valor da qualidade da interseção/triangulação do sistema de visão *stereo*

Capítulo 4

Projeto

4.1	Formulação do Problema	27
4.2	Arquitetura do Sistema	30
4.3	Calibração do Sistema <i>Stereo</i> Ativo	31

4.1 Formulação do Problema

Através de um desafio estimulante do ponto de vista científico, esta dissertação pretende acentuar diversos problemas científicos vocacionados para a área de investigação em Robótica Inteligente e Inteligência Artificial.

Como forma de promover esta iniciativa surgiu o projeto RoboCup, composto por diversas ligas de investigação e competição baseando-se na utilização de um problema standard - o futebol robótico - onde um elevado conjunto de tecnologias é necessário para possibilitar a construção de uma equipa de Robôs reais ou virtuais com o objetivo de participar num jogo de futebol com determinadas regras pré-definidas.

Uma das ligas presentes no RoboCup é a liga de Futebol Robótico Médio (figura 4.1), onde o objetivo principal é o jogo de futebol entre robôs.



Figura 4.1: *Middle Size League* - MSL.

Esta modalidade é uma das que desperta bastante interesse em mais de 20 grupos de investigação espalhados pelo mundo, colocando a competição com um desenvolvimento tecnológico bastante elevado.

O problema endereçado nesta dissertação, está relacionado com a necessidade de melhorar a qualidade da perceção tridimensional dos robôs da equipa de MSL ISePorto, ou seja, com o avanço tecnológico dos dispositivos de chuto, defender os remates dos adversários tornou-se bastante complicado devido à grande velocidade dos remates que são feitos principalmente pelo ar. A equipa ISePorto já tem implementado um algoritmo para estimar a trajetória da bola durante um remate pelo ar [28], mas a fraca perceção da profundidade da bola por vezes resulta em más inicializações da bola, o que pode ser decisivo durante um jogo.

Um dos outros problemas encontrados em cenário de jogo MSL, é a determinação mais precisa da distância dos robôs adversários. Este problema deve-se ao facto de as marcas utilizadas para identificação dos robôs¹ serem pouco visíveis em algumas equipas da competição da liga de Futebol Robótico Médio (figura 4.2 e 4.3), o que torna a sua deteção mais complexa.

¹Na competição MSL são utilizadas duas marcas diferentes para identificar as equipas (Cor magenta e Cor ciano).

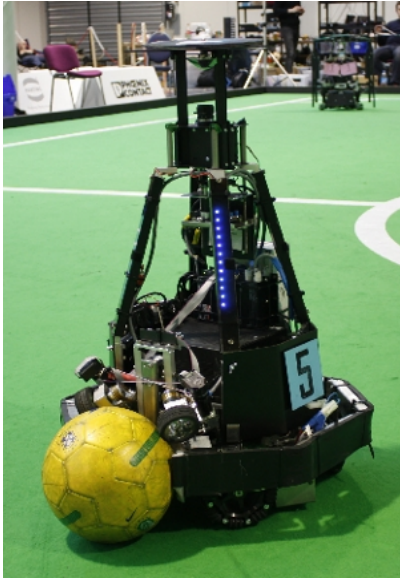


Figura 4.2: Robô da equipa RFC Stuttgart.

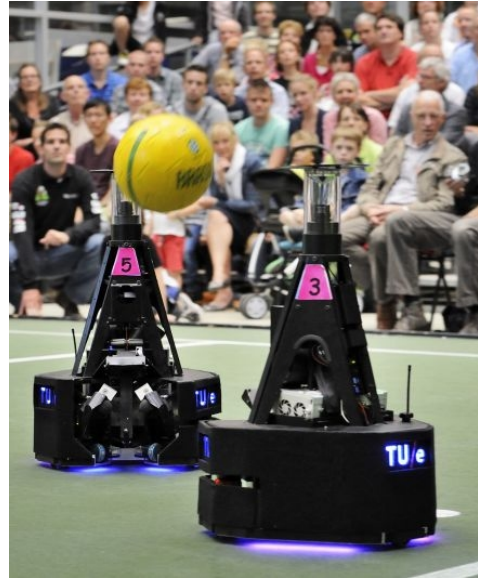


Figura 4.3: Robô da equipa TechUnited.

De forma a melhorar os resultados obtidos, surge a necessidade de estimar a posição 3D da bola através de um sistema de visão *stereo*. Os robôs são constituídos por uma câmara superior que designamos de câmara *head* e por uma câmara inferior que designamos de câmara *kicker*, com exceção do guarda-redes que possui mais uma câmara (*keeper*) para detetar os remates da bola pelo ar (figura 4.4), ou seja, através de duas câmaras podemos criar um sistema de visão *stereo* com o intuito de melhorar a posição do robô aos objetos.

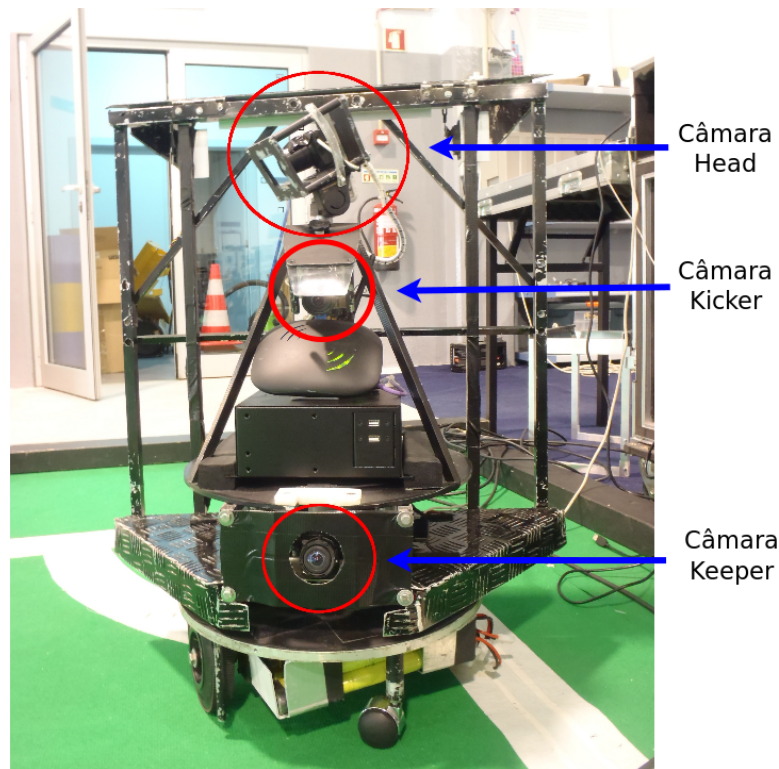


Figura 4.4: Posição das câmaras no robô guarda-redes ISePorto.

No entanto, as câmaras existentes nos robôs, estão montadas em partes móveis (rotativas) o que levanta alguns problemas adicionais ao desenvolvimento de um sistema de visão *stereo*.

Os robôs da equipa MSL do ISePorto, possuem três eixos de rotação (figura 4.5), sendo estes, o eixo de rotação do próprio robô, que designamos de θ_{robot} , o eixo de rotação do chutador para um melhor chute e controle da bola, que designamos de θ_{kicker} e por fim o eixo de rotação da câmara superior, ou seja, câmara da *head* que designamos de θ_{head} .

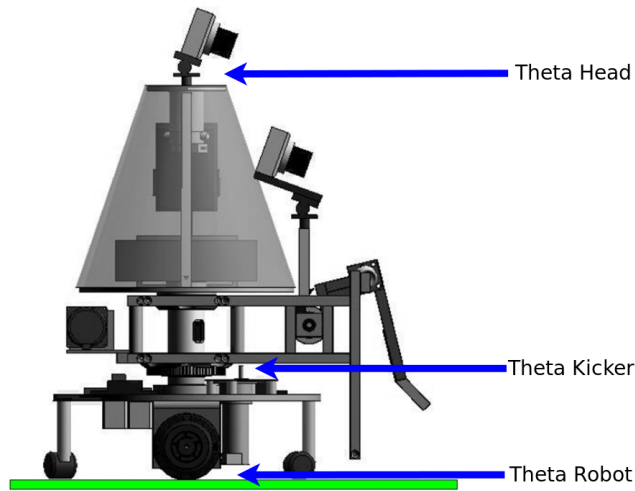


Figura 4.5: Eixos de rotação no robô ISePorto [8].

Devido ao facto do robô ISePorto possuir um eixo de rotação entre a câmara *head* e a câmara *keeper*, foi necessário desenvolver um sistema de visão *stereo* ativo. Assim, através da leitura do sensor *encoder* obtemos o θ_{head} , que associado com uma calibração prévia, e à metodologia que se vai desenvolver neste trabalho, vai permitir a determinação da calibração *stereo* (R, T) de forma dinâmica em função do ângulo lido, sendo possível deste modo determinar a posição 3D dos objetos em determinados ângulos entre as duas câmaras.

4.2 Arquitetura do Sistema

De forma a preencher os requisitos impostos e a implementar um sistema que obtenha os resultados pretendidos foi desenvolvida a arquitetura do sistema conforme apresentado na figura 4.6.

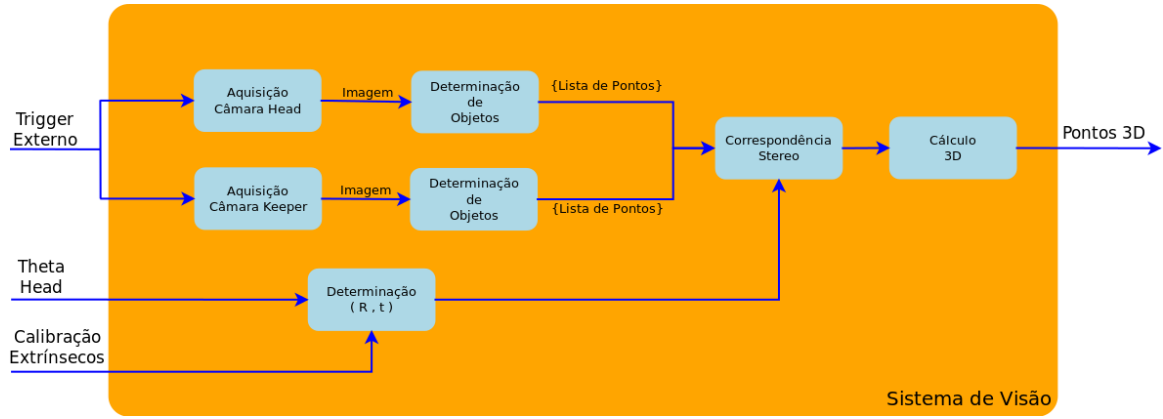


Figura 4.6: Arquitetura sistema de visão e integração no robô IsePorto.

A aquisição de cada câmara é efetuada através de um *trigger* externo de modo a que a captura da imagem de ambas as câmaras seja no mesmo instante, assim como a leitura do ângulo do *encoder*. Após a aquisição das imagens são detetados objetos de interesse pelas duas câmaras, ou seja, a deteção da bola, das marcas dos robôs ou de linhas dos campos e posteriormente elaborada uma lista de pontos de interesse para determinação da sua posição 3D.

De modo a obter a correspondência *stereo* de ambas as imagens é necessário a determinação da matriz de rotação e translação. Como tal, através da leitura do Θ_{Head} do robô e da calibração dos parâmetros extrínsecos de cada câmara é obtido o valor da rotação e translação face ao ângulo de rotação.

Após obtida a lista de pontos de interesse e as matrizes de rotação e translação face ao ângulo de rotação, é efetuada a correspondência *stereo* e a triangulação de forma a obter a posição tridimensional do objeto.

4.3 Calibração do Sistema *Stereo* Ativo

Para determinar a posição 3D de um objeto num sistema de visão *stereo* ativo é necessário obter uma calibração inicial da matriz de rotação e translação de um ângulo previamente conhecido.

Para efetuar a calibração, foi colocado um alvo de xadrez perfeitamente na vertical em frente das duas câmaras e determinada a posição tridimensional de pelo menos 8 pontos conhecidos do alvo ao robô que estava previamente nivelado e na posição $[0,0,0]$ do referencial do *world*. Após determinados os pontos 3D e os pontos projetados no plano da imagem de cada uma das câmaras, podemos determinar a matriz de rotação ($R_{head_{calibrado}}$) e translação ($T_{head_{calibrado}}$) do sistema de visão *stereo* para um dado ângulo de rotação do eixo da câmara

$head$ conhecido ($\theta_{head_{calibrado}}$).



Figura 4.7: Vista do alvo pelas câmaras.

Como forma de validação dos parâmetros extrínsecos de cada câmara partindo do Ponto $[0, 0, 0]$ do referencial da câmara, através da fórmula (4.1), podemos determinar o centro ótico da câmara no referencial do mundo substituindo $P_{CAM} = [0, 0, 0]$ na fórmula (4.1) e resolvendo em ordem a P_{world} como podemos verificar na fórmula (4.2).

$$P_{CAM} = [R|T]P_{world} \quad (4.1)$$

$$P_{CentroOticoWorld} = P_{world} = R^T \times (-T) \quad (4.2)$$

Após a calibração do sistema de visão *stereo* para um dado ângulo θ_{head} conhecido, temos o nosso sistema calibrado obtendo assim a matriz de rotação e translação da calibração.

Uma vez que a calibração inicial foi efetuada de modo aos eixos de rotação do robô coincidir com o eixo dos Z do mundo, para determinar a matriz de rotação para um novo ângulo θ_{head} , precisamos de multiplicar a matriz de rotação da calibração por uma matriz de rotação segundo o eixo dos zz através da fórmula (4.5).

$$\Delta\theta_{head} = \theta_{head_{new}} - \theta_{head_{calibrado}} \quad (4.3)$$

$$R_z = \begin{bmatrix} \cos(\Delta\theta_{head}) & \sin(\Delta\theta_{head}) & 0 \\ -\sin(\Delta\theta_{head}) & \cos(\Delta\theta_{head}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

$$R_{head_{new}} = R_z \times R_{head_{calibrado}} \quad (4.5)$$

A matriz de translação para um novo ângulo, pode ser obtida através da seguinte fórmula (4.7), ou seja, se multiplicarmos a matriz de rotação da calibração transposta pela matriz de translação negada (4.6), obtemos o Ponto $[0, 0, 0]$ do referencial da câmara *head* da calibração. Por sua vez, após determinado o Ponto $[0, 0, 0]$ do referencial da câmara, aplicando a fórmula (4.7) obtemos a matriz de translação do ângulo pretendido.

$$T_{robot} = R_{head_{calibrado}}^t \times (-T_{head_{calibrado}}) \quad (4.6)$$

$$T_{head_{new}} = -R_{head_{new}}(R_z \times T_{robot}) \quad (4.7)$$

Após obter os novos valores da rotação e translação da câmara *head*, podemos calcular a posição de um Ponto 3D no mundo através do cálculo da triangulação.

Esta página foi intencionalmente deixada em branco.

Capítulo 5

Implementação

5.1	Implementação do sistema de aquisição	35
5.1.1	Câmaras CCD	35
5.1.2	Dispositivo de disparo das câmaras e leitura do ângulo de rotação . .	36
5.2	Implementação do Sistema de Visão <i>Stereo</i> Activo	37
5.2.1	Geometria epipolar	38
5.2.2	Parâmetros extrínsecos e calibração <i>stereo</i> ativa	38
5.2.3	Determinação de objetos nas imagens	39
5.2.4	Determinação do Ponto 3D no mundo	40

5.1 Implementação do sistema de aquisição

Para o desenvolvimento da presente dissertação, foram utilizados os robôs da equipa ISePorto. Estes robôs são compostos por uma plataforma sensorial no qual no decorrer deste capítulo são expostas as características fundamentais de cada sensor e posteriormente é estabelecida a disposição dos mesmos nos robôs.

5.1.1 Câmaras CCD

O sistema inclui duas câmaras CCD do fabricante UeYE, modelo UI-1225-LE. A resolução máxima é de $752(h) \times 480(v)$ com *frame rate* de $30frames/s$. Estas câmaras tem à disposição pinos GPIO que, entre outras funções, permitem o disparo controlado por dispositivos externos, ou seja, através de um *trigger* externo, a câmara faz aquisição de um *frame*. A comuni-

cação utilizada na transferência de informações das câmaras para a plataforma computacional respeita o protocolo USB 2.0.



Figura 5.1: Câmera UeYE modelo UI-1225-LE [9].

5.1.2 Dispositivo de disparo das câmaras e leitura do ângulo de rotação

Para efetuar o disparo sincronizado das duas câmaras utilizadas no sistema de visão *stereo*, foi desenvolvido no LSA uma placa de *trigger* externo, sendo controlado o disparo pela porta RS-232 no pino RTS da plataforma computacional. Esta placa utiliza foto-acoplador para proteção da câmara.



Figura 5.2: Dispositivo de disparo das câmaras.

Para determinação da rotação e do ângulo da câmara superior do robô, existe um *encoder* ótico (figura 5.3) que é lido por um sistema embebido designado de Pandora (figura 5.4), que por sua vez comunica com a plataforma computacional.

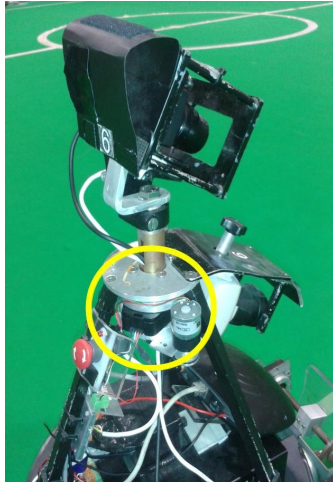


Figura 5.3: Encoder para determinação do ângulo da câmara superior.

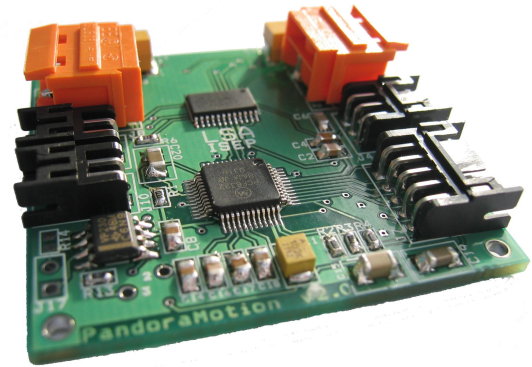
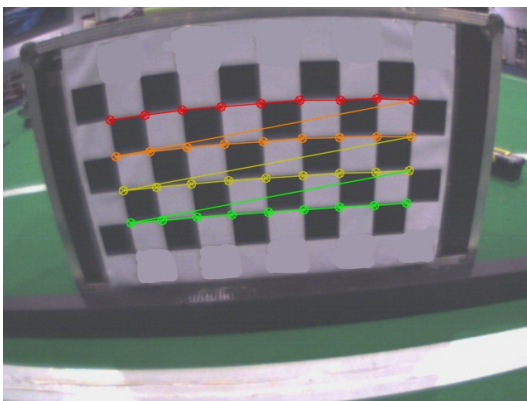


Figura 5.4: Sistema Embebido Pandora Motion [8].

5.2 Implementação do Sistema de Visão *Stereo* Activo

Para a implementação do Sistema de Visão *stereo*, foi utilizado o OpenCV (*Open Source Computer Vision*) que é uma biblioteca de funções relacionadas com visão por computador, como por exemplo funções de processamento de imagem, detecção, reconhecimento, monitorização, entre outras.

Para efetuar a calibração do Sistema de Visão, foi utilizado um alvo colocado em frente do robô de modo a que seja visível por ambas as câmaras e seguido o procedimento referido no capítulo anterior. Para reconhecer os cantos internos do alvo, foi utilizada a função *cvFindChessboardCorners* da biblioteca OpenCV, recebendo a imagem adquirida por cada câmara e as dimensões do tabuleiro. Esta função retorna o número de cantos encontrados e a posição de cada um [29].



(a) Head



(b) Keeper

Figura 5.5: Detecção dos cantos encontrados no alvo.

5.2.1 Geometria epipolar

Através da geometria epipolar utilizando a matriz essencial E (5.1), é possível calcular a linha epipolar numa imagem associada a um ponto da outra imagem. Utilizando a função *cvComputeCorrespondEpilines* obtemos a equação geral da reta no formato $ax + by + c = 0$.

$$E = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} R \quad (5.1)$$

Para desenhar as linhas epipolares na imagem, necessitamos de aplicar a equação da reta no formato $y = mx + b$. O valor de m e b é obtido através de:

$$m = -\frac{b}{a} \quad (5.2)$$

$$b = -\frac{c}{a} \quad (5.3)$$

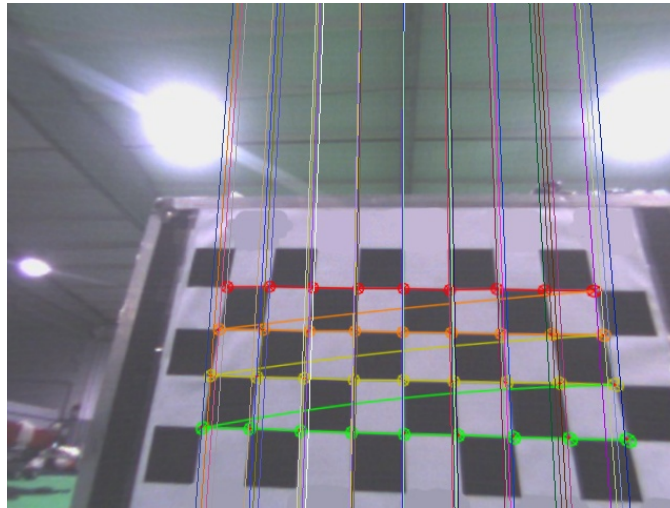


Figura 5.6: Linhas Epipolares.

5.2.2 Parâmetros extrínsecos e calibração *stereo* ativa

Para determinar os parâmetros extrínsecos da câmara (matriz de rotação e translação), foi utilizada a função *cvFindExtrinsicCameraParams2* que recebe como parâmetros os pontos no espaço do mundo, os pontos no espaço da imagem e o número de cantos do *frame*. Além disso, recebe a matriz de intrínsecos e os coeficientes de distorção da câmara obtidos através da *ToolBox Camera Calibration* para *Matlab*. O vetor de translação e o de rotação, são obtidos na forma de ângulos de Rodrigues. Para transformar o vetor de rotação numa matriz de rotação 3x3 convencional é utilizada a função *cvRodrigues2*, obtendo assim a matriz dos

parâmetros extrínsecos (5.4) para cada uma das câmaras.

$$Ext = \begin{bmatrix} R_1 & R_2 & R_3 & t_x \\ R_4 & R_5 & R_6 & t_y \\ R_7 & R_8 & R_9 & t_z \end{bmatrix} \quad (5.4)$$

Após a determinação dos parâmetros extrínsecos obtemos as matrizes de rotação e translação da câmara *head*, R_h e T_h , e da câmara *keeper*, R_k e T_k . Existe uma relação da matriz de rotação R e da translação T , entre as duas câmaras, relação esta que pode ser obtida pelas seguintes fórmulas (5.5) e (5.6).

$$R = R_k \times (R_h)^T \quad (5.5)$$

$$T = T_k - R \times T_h \quad (5.6)$$

O processo de calibração do sistema *stereo* está concluído, obtendo assim os extrínsecos de calibração da câmara *keeper* ($Ext_{keeper-calib}$) e os extrínsecos da câmara *head* ($Ext_{head-calib}$) para um dado ângulo ($\Theta_{head-calib}$) conhecido.

Sendo as matrizes R_h e T_h das equações (5.5) e (5.6), determinamos dinamicamente em função do ângulo do *encoder head* no instante do *trigger* os parâmetros extrínsecos calibrados da câmara *head*.

5.2.3 Determinação de objetos nas imagens

Os robôs da equipa de futebol ISePorto estão embebidos de um método de segmentação de cor que permite uma rápida análise do meio ambiente que os rodeia. Como o cenário de jogo é em ambientes interiores, a segmentação de cor é o método utilizado para estruturar a imagem, de forma a obter uma correta identificação dos objetos da imagem [10].

Como tal, o sistema de segmentação de cor junta a informação em *blobs*¹, permitindo agregar rapidamente a informação dos diferentes pixels em regiões comuns. Estas regiões são compostas por determinados atributos, nomeadamente:

- Área da imagem em coordenadas 2D da imagem (x_i, y_i, x_f, y_f) .
- Centro de massa em coordendas 2D da imagem (x_{cm}, y_{cm}) .
- Cor

¹Os blobs, permitem juntar rapidamente a informação dos diferentes pixels em regiões comuns, sendo agregados obedecendo a um critério ou propriedade que as distinga entre si.

Com estes dados, é possível extrair a informação dos objetos de interesse, obtendo assim as coordenadas dos pontos vistas no plano da imagem por cada câmara.

5.2.4 Determinação do Ponto 3D no mundo

Para determinar a posição 3D de um ponto no mundo, é determinado o valor de $\Delta\theta_{head}$ (4.3) para de seguida determinar o valor da matriz de rotação R (4.5) e a matriz de translação T (4.7).

Após obter os novos valores da rotação e translação da câmara *head*, podemos calcular a posição de um Ponto 3D no mundo através da função *cvTriangulatePoints* que recebe como parâmetros os valores extrínsecos de ambas as câmaras e a coordenada dos pontos vistas no plano da imagem por cada câmara.

Capítulo 6

Resultados

6.1	Determinação da posição da bola	41
6.1.1	Posição da bola com a câmara <i>head</i> e <i>keeper</i>	42
6.1.2	Posição da bola com a câmara <i>head</i> e <i>kicker</i>	44
6.2	Determinação da distância aos robôs adversários	45
6.3	Determinação da distância as linhas de campo	46
6.4	Validação da visão <i>stereo</i> ativo	46

Neste capítulo são apresentados resultados obtidos como forma de validação da arquitetura do sistema, da metodologia de determinação dinâmica da calibração *stereo* e do sistema implementado. Foram efetuados diversos testes em cenário de jogo MSL, entre eles a determinação da distância da bola ao robô, para diferentes distâncias e alturas, bem como para diferentes ângulos da câmara *head*, a posição de robôs adversários e distância as linhas de limitação do campo de futebol.

6.1 Determinação da posição da bola

Como forma de validação dos resultados obtidos da posição 3D da bola retornada pelo sistema de visão *stereo*, foram efetuadas medições da bola ao robô através de fita métrica como forma de *ground truth* e foram comparados os valores obtidos entre duas *baselines* diferentes (visão *stereo* com a câmara *head* e *keeper* e visão *stereo* com a câmara *head* e *kicker*).

6.1.1 Posição da bola com a câmara *head* e *keeper*

Nas figuras 6.1 e 6.2 podemos observar a deteção da bola pelas duas câmaras em situações diferentes. Na situação 1 estamos perante uma bola junto ao solo e na situação 2 uma bola aérea. A comparação dos resultados obtidos pelo sistema de visão *stereo* com os valores exatos estão expressos na tabela 6.5.



(a) Câmara Head



(b) Câmara Keeper

Figura 6.1: Determinação da posição da bola na situação 1.



(a) Câmara Head



(b) Câmara Keeper

Figura 6.2: Determinação da posição da bola na situação 2.

Tabela 6.1: Resultados obtidos da posição da bola entre a câmara *head* e *keeper*

	Medida exata ¹			Medida obtida pelo sistema de visão <i>stereo</i>		
	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
Situação 1	277 <i>cm</i>	0 <i>cm</i>	11 <i>cm</i>	268,9 <i>cm</i>	0,1 <i>cm</i>	14,6 <i>cm</i>
Situação 2	349 <i>cm</i>	0 <i>cm</i>	54 <i>cm</i>	336,0 <i>cm</i>	-2,5 <i>cm</i>	56,3 <i>cm</i>

¹Medições efetuadas com auxílio da fita métrica em *cm*.

Como se pode constatar, os erros do sistema de visão *stereo* são inferiores a 4% para uma distância de aproximadamente 3,5 m e com a bola a 56 cm de altura. Este valor é uma melhoria significativa face à implementação atual nos robôs, onde a distância às bolas que não estavam assentes sobre o plano do campo é obtida pela dimensão do objeto introduzindo elevados erros na medição da distância.

Como forma de validar a distância máxima detetável de uma bola aérea pelo sistema de visão ativo, foram efetuadas diversas medições até uma distância máxima de 5 metros.

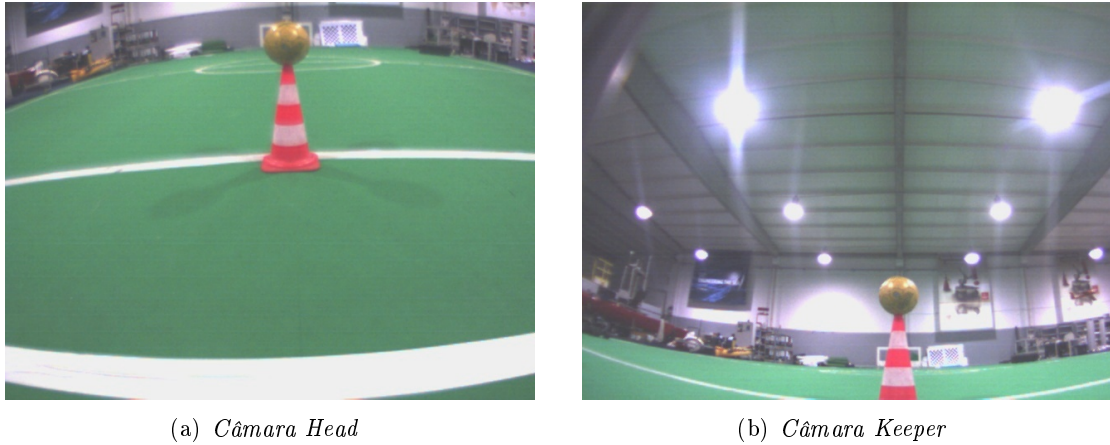


Figura 6.3: Determinação da posição 3D de uma bola aérea.

Tabela 6.2: Resultados obtidos da posição da bola aérea entre a câmara *head* e *keeper*

	Medida exata ¹			Medida obtida pelo sistema de visão <i>stereo</i>		
	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
Bola 1 m	100 cm	0 cm	64 cm	95,9 cm	0,4 cm	61,6 cm
Bola 1,5 m	150 cm	0 cm	64 cm	137,0 cm	-1,3 cm	62,3 cm
Bola 2 m	200 cm	0 cm	64 cm	189,3 cm	-2,4 cm	63,2 cm
Bola 2,5 m	250 cm	0 cm	64 cm	239,1 cm	2,1 cm	64,5 cm
Bola 3 m	300 cm	0 cm	64 cm	278,7 cm	-1,5 cm	65,3 cm
Bola 3,5 m	350 cm	0 cm	64 cm	335,8 cm	-1,1 cm	65,7 cm
Bola 4 m	400 cm	0 cm	64 cm	367,0 cm	-1,5 cm	64,5 cm
Bola 4,5 m	450 cm	0 cm	64 cm	410,0 cm	-3,1 cm	63,1 cm
Bola 5 m	500 cm	0 cm	64 cm	446,0 cm	-4,4 cm	61,3 cm

Através da tabela 6.2, podemos observar que a posição da bola até 4,5m tem um erro sempre inferior a 10%, só ultrapassando ligeiramente este valor para 5m.

6.1.2 Posição da bola com a câmara *head* e *kicker*

Os dados apresentados na tabela 6.3 apresentam a determinação da posição da bola para dois ângulos diferentes.



(a) Câmera Head



(b) Câmera Kicker

Figura 6.4: Determinação da posição da bola na situação 3.



(a) Câmera Head



(b) Câmera Kicker

Figura 6.5: Determinação da posição da bola na situação 4.

Tabela 6.3: Resultados obtidos da posição da bola entre a câmara *head* e *kicker*

	Medida exata ¹			Medida obtida pelo sistema de visão <i>stereo</i>			Θ_{head}
	X	Y	Z	X	Y	Z	
Situação 3	212 cm	0 cm	11 cm	206,3 cm	1,4 cm	10,3 cm	0°
Situação 4	127 cm	-41 cm	11 cm	124,2 cm	-39,3 cm	13,6 cm	-22,2°

Podemos observar que mesmo para uma variação elevada do ângulo da câmara *head* os erros mantêm a mesma ordem de grandeza.

¹Medições efetuadas com auxílio da fita métrica em *cm*.

6.2 Determinação da distância aos robôs adversários

Neste conjunto de teste é avaliada a qualidade das medições *stereo* para outro tipo de objetos, mais concretamente robô. Para este tipo de objetos, são utilizados os centros de massa dos *blobs* das marcas, neste caso, as marcas magenta. É utilizado o robô guarda-redes que é o único que possui 3 câmaras, e compara-se a qualidade das medidas *stereo*, para duas combinações de pares de câmaras, sendo elas Câmara *Head* + Câmara *Keeper* e Câmara *Head* + Câmara *Kicker*. São avaliadas duas situações, com ângulo da *head* próximo de zero e com um ângulo da *head* de $22,5^\circ$.

Na tabela 6.4 podemos comparar as distâncias obtidas do robô adversário através da visão *stereo* com a câmara *Head* e *Kicker* e com a visão *stereo* com a câmara *Head* e *Keeper*.

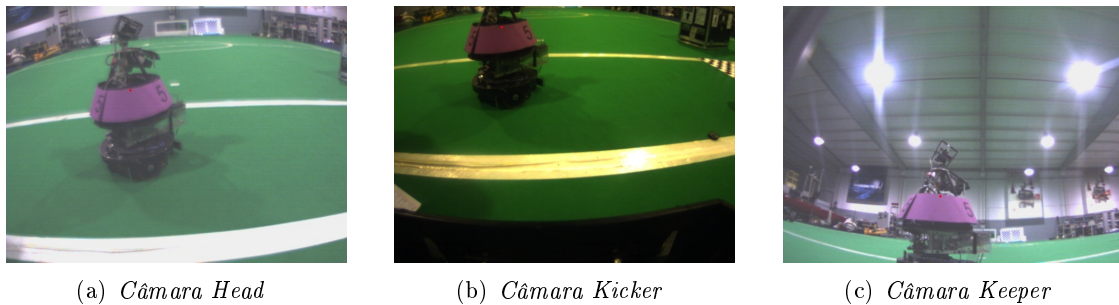


Figura 6.6: Determinação da posição do robô adversário na situação 1.

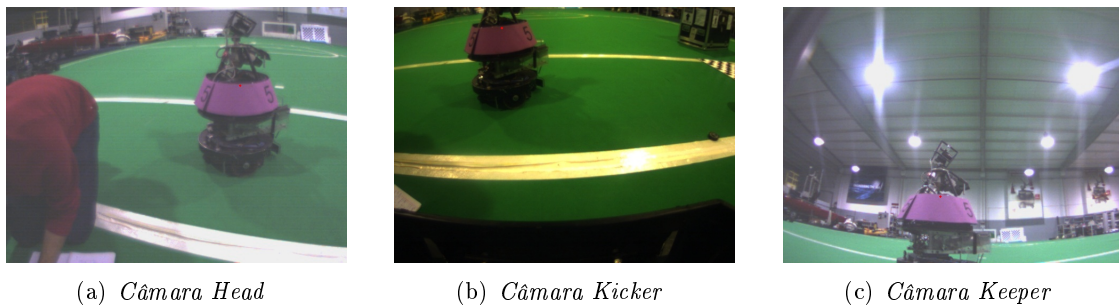


Figura 6.7: Determinação da posição do robô adversário na situação 2.

Tabela 6.4: Comparação da distância ao robô adversário com duas *baselines* diferentes.

	Medida exata ¹			Stereo Head e Kicker			Stereo Head e Keeper			Θ_{head}
	X	Y	Z	X	Y	Z	X	Y	Z	
Situação 1	106 cm	20 cm	43 cm	110,3 cm	17,3 cm	41,2 cm	113,0 cm	18,4 cm	42,1 cm	$-0,5^\circ$
Situação 2	106 cm	20 cm	43 cm	110,8 cm	14,9 cm	40,7 cm	115,3 cm	17,9 cm	41,7 cm	$22,5^\circ$

Dos resultados obtidos, podemos constatar que o sistema obtém resultados idênticos independentemente do ângulo da câmara *head*.

¹Medições efetuadas com auxílio da fita métrica em cm.

6.3 Determinação da distância as linhas de campo

Seguidamente, vamos avaliar o funcionamento do sistema *stereo* desenvolvido no que diz respeito, a medições do posicionamento de linhas do campo.

Na figura 6.8 que se segue podemos observar o canto da grande área do campo de futebol MSL. Na tabela 6.4 podemos comparar os resultados obtidos com as medidas *ground truth* obtidas com auxílio da fita métrica.

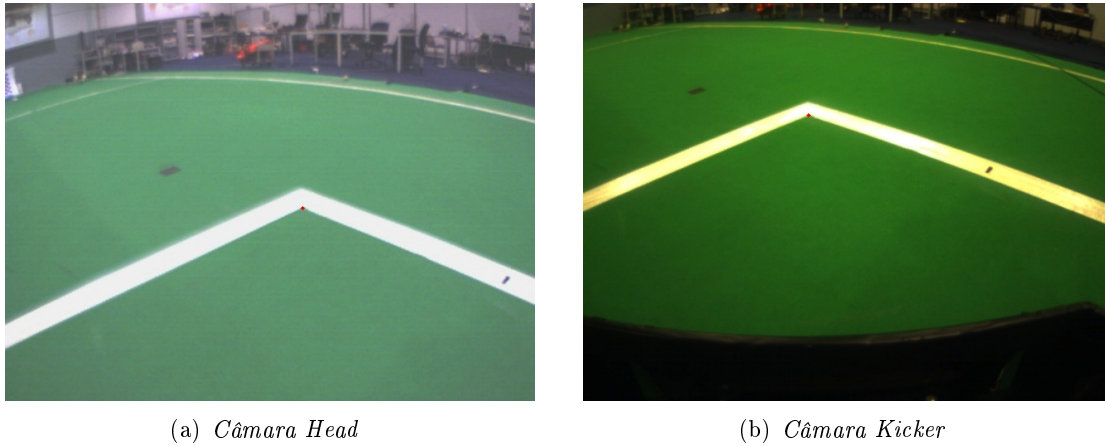


Figura 6.8: Determinação da distância do canto ao robô.

Tabela 6.5: Resultados obtidos da posição da bola entre a câmara *head* e *keeper*

Medida exata ¹			Medida obtida pelo sistema de visão <i>stereo</i>			Θ_{Head}
X	Y	Z	X	Y	Z	
132 cm	0 cm	0 cm	127,3 cm	1,6 cm	0,8 cm	8,5°

6.4 Validação da visão *stereo* ativo

Como forma de validação do sistema de visão *stereo* ativo em situações de movimento da câmara *head* em tempo real, foram efetuadas diversas medições à bola de diferentes ângulos do Θ_{Head} . Este teste serviu para analisar o comportamento do sistema em diversos ângulos. Na figura 6.3 podemos observar o posicionamento da bola face à câmara *Head* e *Keeper*. Nos gráficos seguintes podemos comparar a posição 3D da bola relativamente a um ângulo de variação entre -25° e 25° . As medidas de *ground truth* foram efetuadas com auxílio da fita métrica sendo os valores $x=235cm$ $y=0cm$ $z=64cm$.

¹Medições efetuadas com auxílio da fita métrica em *cm*.

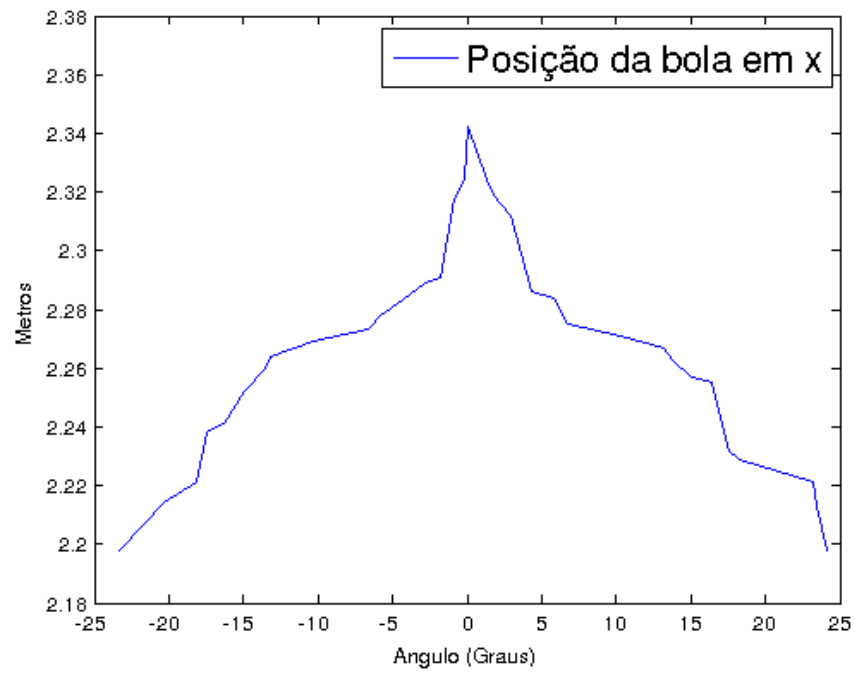


Figura 6.9: Posição da bola em X relativamente ao valor do eixo de rotação da câmara *head*.

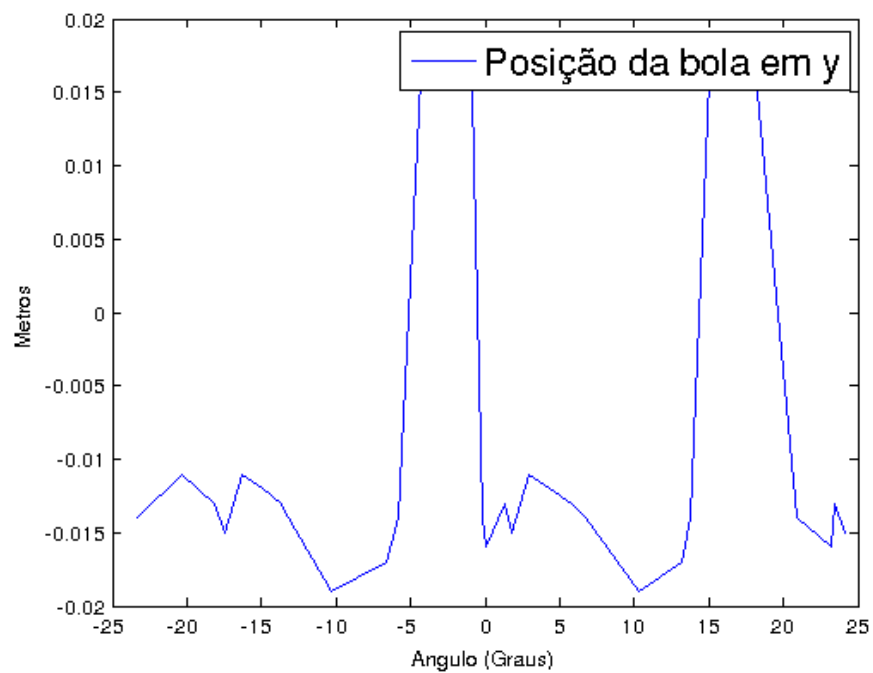


Figura 6.10: Posição da bola em Y relativamente ao valor do eixo de rotação da câmara *head*.

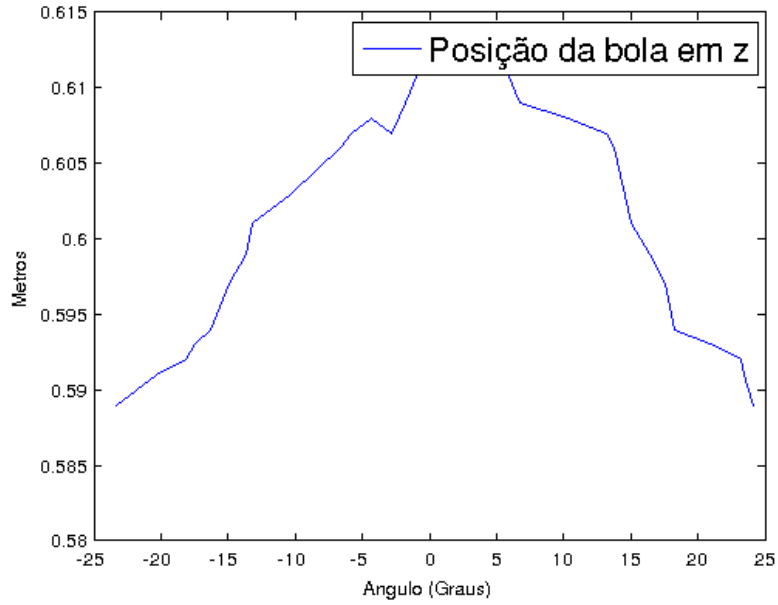


Figura 6.11: Posição da bola em Z relativamente ao valor do eixo de rotação da câmara *head*.

Os gráficos das figuras 6.9, 6.10 e 6.11 apresentam os resultados obtidos da distância da bola para determinados ângulos possíveis do eixo de rotação da câmara *head*. Contudo, podemos observar nos gráficos das figuras 6.12 e 6.13 que apesar do erro do nosso sistema aumentar com o afastamento angular relativamente ao ângulo de calibração, o valor percentual do erro atinge nos ângulos extremos cerca de 6%. Sendo inferior a 3% num intervalo de cerca de 30 segundos.

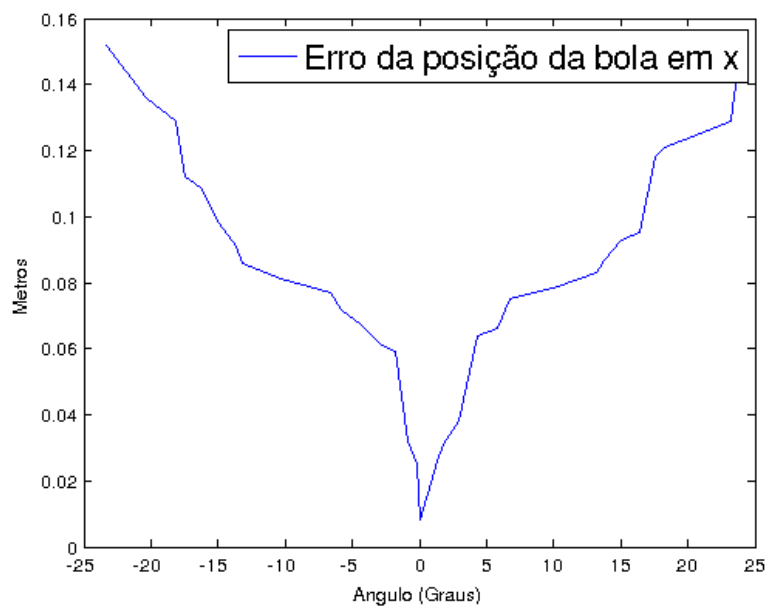


Figura 6.12: Erro da posição da bola em X relativamente ao valor da medida exata.

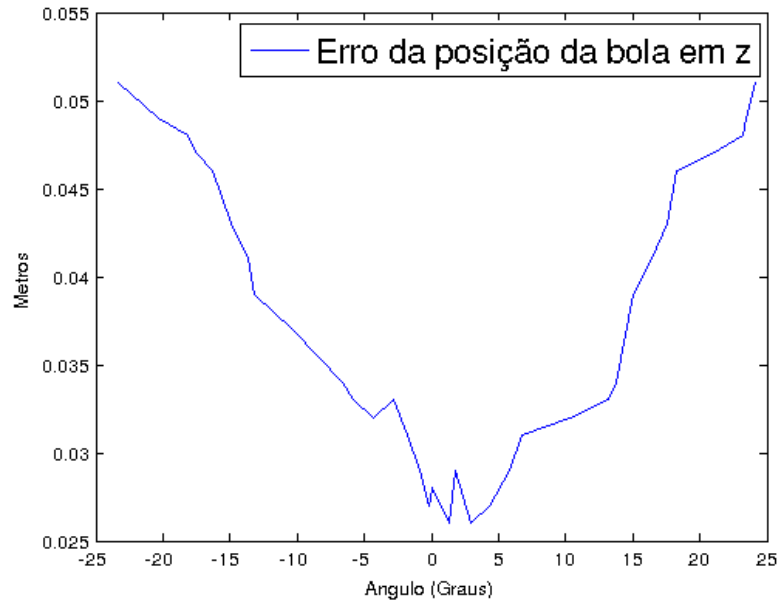


Figura 6.13: Erro da posição da bola em Z relativamente ao valor da medida exata.

O ligeiro aumento do erro com o afastamento angular da posição de calibração, leva-nos a concluir que, ou o eixo de rotação do robô não está mecanicamente na vertical face ao referencial do robô, ou existiu um pequeno erro no *setup* de calibração, ou seja, o alvo ou o robô não estavam completamente alinhados. Deste modo, perspectiva-se que seja possível melhorar a calibração, ou através de um *setup* mais exato, ou através de uma melhoria no método que permita a estimação deste desvio.

Como forma de melhorar a qualidade dos resultados obtidos, existe a necessidade de melhorar o sistema de sincronismo entre o ângulo de rotação e o *trigger* das câmaras, possibilitando assim testes mais exaustivos com objetos em movimento e o próprio robô em movimento.

Esta página foi intencionalmente deixada em branco.

Capítulo 7

Conclusão e Trabalho Futuro

O trabalho abordou o desenvolvimento de um sistema de visão *stereo* ativo para os robôs de futebol robótico da equipa ISePorto do ISEP. Este trabalho foi motivado pela necessidade de melhorar as capacidades de perceção do ambiente por parte dos robôs, imposta pelo aumento da dinâmica que se tem vindo a verificar nas competições do RoboCup, a nível de velocidade dos robôs e aos remates mais fortes e aéreos.

A atual utilização de visão monocular apresenta sérias limitações em termos de perceção tridimensional, ou seja, a aplicação de abordagens *stereo* tradicionais nos robôs, estão limitadas pelo facto das câmaras existentes nos robôs estarem montadas em eixos rotativos.

Para endereçar este problema, nesta tese foi projetada uma arquitetura para um sistema de visão *stereo* ativo, que utilize a informação proveniente das diversas câmaras rotativas para perceção tridimensional.

Foi implementado um módulo para determinação da calibração *stereo* em função das medidas de ângulos dos vários eixos.

Foi desenvolvido e integrado um mecanismo de disparo para a aquisição sincronizada das imagens de várias câmaras e leitura do ângulo dos eixos de rotação.

Foram implementados módulos para a determinação de medidas tridimensionais a diversos objetos tais como bolas, robôs e marcações do campo.

O sistema de visão *stereo* ativo desenvolvido foi testado e validado experimentalmente, tendo sido efetuados diversos testes a nível de deteção de bolas aéreas e bolas rente ao chão, a deteção das marcas dos robôs adversários e a deteção de linhas do campo a diversas distâncias e a diversos ângulos do eixo de rotação da câmara superior.

Os resultados obtidos apresentam uma melhoria significativa, na perceção tridimensional

dos objetos, a nível da deteção de bolas e robôs adversários relativamente a visão monocular utilizada anteriormente.

Os testes efetuados na determinação da posição tridimensional da bola para um distância aproximadamente de $3,5m$ e com a bola a $56m$ de altura, os erros obtidos do sistema de visão *stereo* são inferiores a 4%, sendo notável uma melhoria significativa na determinação da distância face às bolas que não estavam assentes sobre o plano do campo.

Podemos observar nos resultados obtidos que mesmo para uma variação elevada do ângulo da câmara *head* os erros mantêm a mesma ordem de grandeza.

Em termos de trabalho futuro perspectiva-se a validação do sistema de visão *stereo* ativo em cenário de jogo, validando situações de deteção de vários objetos no mesmo instante, de forma a existir a correspondência correta entre eles.

Considero também desenvolver um sistema de visão *ground truth* externo como forma de comparação da posição tridimensional obtida pelo robô com a obtida pelo sistema de visão externo.

Bibliografia

- [1] M. Shibata and T. Honma, “3D object tracking on active stereo vision robot,” in *Advanced Motion Control, 2002. 7th International Workshop on*, 2002, pp. 567–572.
- [2] Y. Nakabo, T. Mukai, Y. Hattori, Y. Takeuchi, and N. Ohnishi, “Variable Baseline Stereo Tracking Vision System Using High-Speed Linear Slider,” in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005, pp. 1567–1572.
- [3] W. Klarquist and A. Bovik, “Adaptive variable baseline stereo for vergence control,” in *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, vol. 3, 1997, pp. 1952 –1959 vol.3.
- [4] “Modelo Câmara Pinhole.” [Online]. Available: <http://cga-epatv.blogspot.pt/2010/09/pinhole.html>
- [5] R. I. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision.” Cambridge University Press, 2004.
- [6] “Stereo Rectification,” in http://fit.com.ru/Projects/stereo_rectification.htm *Ultimo acesso Outubro 2012*.
- [7] Emanuele Trucco and Alessandro Verri, *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [8] A. P. Dias, “Sistema de Controlo Embebido para Locomoção de Sistemas Autónomos.” Master thesis: Instituto Superior Técnico, 2007.
- [9] “USB2 UeYE Camera.” [Online]. Available: <http://www.ids-imaging.com>
- [10] H. M. Silva, “Sistema de Visão em Tempo Real para Sistemas Autónomos.” Master thesis: Instituto Superior Técnico, 2007.

- [11] D. Marr, "A Computational Investigation into the Human Representation and Processing of Visual Information," Ph.D. dissertation, W. H. Freeman & Co., San Francisco, USA, 1982.
- [12] S. T. Bernard, "A stochastic approach to stereo vision," Philadelphia, USA, 1986.
- [13] R. S. Nourbakhsh and I. R., "Introduction to Autonomous Mobile Robots," in *Bradford Company, Scituate, MA, USA*, 2004.
- [14] R. M. H.D. Burkhard, D. Duhaut, M. Fujita, P. Lima and R. Rojas, "The road to robocup 2050," *Robotics & Automation Magazine, IEEE*, 9(2):31-38.
- [15] M. Shibata and N. Kobayashi, "Image-based visual tracking for moving targets with active stereo vision robot," in *SICE-ICASE, 2006. International Joint Conference*, 2006, pp. 5329–5334.
- [16] J. K. Yi Ma, Stefano Soatto and S. S. Sastry, "An Invitation to 3-D Vision, from images to geometric models." in *Springer*, 2006.
- [17] D. G. B. Kaehler. and D. Adrian, "LearningOpenCV." O'Reilly, 2008.
- [18] J. Y. Bouguet, "Camera Calibration Toolbox for Matlab." [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/parameters.html
- [19] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction." in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 1106–1112.
- [20] Y. H. Lee, J. S. & Jeong, "CCD camera calibration and projection error analysis," in *4th Korea-Russia International Symposium on Science and Technology, vol.2, no.27, pp. 50-55*, 2000.
- [21] B. Jahne, *Computer Vision and Applications A Guide for Students and Practitioners*.
- [22] J. Bigun, *Vision with Direction A Systematic Introduction to Image Processing and Computer Vision*.
- [23] P. M. David Gallup, Jan-Michael Frahm and M. Pollefeys, "Variable baseline/resolution stereo," 2008.
- [24] S. M. Nixon, *Feature Extraction in Computer Vision and Image Processing*.

- [25] R. I. Hartley., “Theory and Practice of Projective Rectification.” *Internacional Journal of Computer Vision*, 35(2): 115-127, 1999.
- [26] E. T. Andrea Fusiello and A. Verri, “A compact algorithm for rectification of stereo pairs.” in *Machine Vision and Applications*, 12(1): 16-22, p. 2000.
- [27] C. . Z. Z. Loop, “Computing Rectifying Homographies for Stereo Vision,” in *IEEE Conference on Computer Vision and Pattern Recognition, Colorado, USA*, 1999.
- [28] J. Almeida, A. Martins, and et al, “ISePorto Robotic Soccer Team for Robocup 2011: Improving Defence and Dynamic Passing,” Instituto Superior de Engenharia do Porto, 2011.
- [29] “OpenCV. (Open Source Computer Vision Library).” [Online]. Available: www.opencv.org

